



This is a digital copy of a book that was preserved for generations on library shelves before it was carefully scanned by Google as part of a project to make the world's books discoverable online.

It has survived long enough for the copyright to expire and the book to enter the public domain. A public domain book is one that was never subject to copyright or whose legal copyright term has expired. Whether a book is in the public domain may vary country to country. Public domain books are our gateways to the past, representing a wealth of history, culture and knowledge that's often difficult to discover.

Marks, notations and other marginalia present in the original volume will appear in this file - a reminder of this book's long journey from the publisher to a library and finally to you.

### Usage guidelines

Google is proud to partner with libraries to digitize public domain materials and make them widely accessible. Public domain books belong to the public and we are merely their custodians. Nevertheless, this work is expensive, so in order to keep providing this resource, we have taken steps to prevent abuse by commercial parties, including placing technical restrictions on automated querying.

We also ask that you:

- + *Make non-commercial use of the files* We designed Google Book Search for use by individuals, and we request that you use these files for personal, non-commercial purposes.
- + *Refrain from automated querying* Do not send automated queries of any sort to Google's system: If you are conducting research on machine translation, optical character recognition or other areas where access to a large amount of text is helpful, please contact us. We encourage the use of public domain materials for these purposes and may be able to help.
- + *Maintain attribution* The Google "watermark" you see on each file is essential for informing people about this project and helping them find additional materials through Google Book Search. Please do not remove it.
- + *Keep it legal* Whatever your use, remember that you are responsible for ensuring that what you are doing is legal. Do not assume that just because we believe a book is in the public domain for users in the United States, that the work is also in the public domain for users in other countries. Whether a book is still in copyright varies from country to country, and we can't offer guidance on whether any specific use of any specific book is allowed. Please do not assume that a book's appearance in Google Book Search means it can be used in any manner anywhere in the world. Copyright infringement liability can be quite severe.

### About Google Book Search

Google's mission is to organize the world's information and to make it universally accessible and useful. Google Book Search helps readers discover the world's books while helping authors and publishers reach new audiences. You can search through the full text of this book on the web at <http://books.google.com/>

C

562,822

THE DECENNIAL PUBLICATIONS OF  
THE UNIVERSITY OF CHICAGO

PHYSICS    CHEMISTRY    GEOLOGY  
MATHEMATICS

---

---

*GENERAL LIBRARY of the*  
*UNIVERSITY OF MICHIGAN*

---

---

—PRESENTED BY—

*University of Chicago*

*Sept 23 1904*

AS  
36  
.C53d



**THE DECENNIAL PUBLICATIONS OF  
THE UNIVERSITY OF CHICAGO**

# **THE DECENNIAL PUBLICATIONS**

**ISSUED IN COMMEMORATION OF THE COMPLETION OF THE FIRST TEN  
YEARS OF THE UNIVERSITY'S EXISTENCE**

**AUTHORIZED BY THE BOARD OF TRUSTEES ON THE RECOMMENDATION  
OF THE PRESIDENT AND SENATE**

**EDITED BY A COMMITTEE APPOINTED BY THE SENATE**

**EDWARD CAPPS**

**STARR WILLARD CUTTING**

**ROLLIN D. SALISBURY**

**JAMES ROWLAND ANGELL**

**WILLIAM I. THOMAS**

**SHAILER MATHEWS**

**CARL DARLING BUCK**

**FREDERIC IVES CARPENTER**

**OSKAR BOLZA**

**JULIUS STIEGLITZ**

**JACQUES LOEB**

**THESE VOLUMES ARE DEDICATED**  
**TO THE MEN AND WOMEN**  
**OF OUR TIME AND COUNTRY WHO BY WISE AND GENEROUS GIVING**  
**HAVE ENCOURAGED THE SEARCH AFTER TRUTH**  
**IN ALL DEPARTMENTS OF KNOWLEDGE**





## **INVESTIGATIONS**



THE UNIVERSITY OF CHICAGO, *University*  
FOUNDED BY JOHN D. ROCKEFELLER

# INVESTIGATIONS REPRESENTING THE DEPARTMENTS

PHYSICS CHEMISTRY GEOLOGY  
MATHEMATICS

THE DECENNIAL PUBLICATIONS  
FIRST SERIES VOLUME IX

CHICAGO  
THE UNIVERSITY OF CHICAGO PRESS  
1904

*Copyright 1903*  
**BY THE UNIVERSITY OF CHICAGO**

## CONTENTS

I. THE VELOCITY OF LIGHT - - - - -	1
By ALBERT A. MICHELSON, Professor and Head of the Department of Physics	
II. CONCERNING THE GEODESIC CURVATURE AND THE ISOPERIMETRIC PROBLEM ON A GIVEN SURFACE - - - - -	11
By OSKAR BOLZA, Professor of Mathematics	
III. PROOF OF THE SUFFICIENCY OF JACOBI'S CONDITION FOR A PERMANENT SIGN OF THE SECOND VARIATION IN THE SO-CALLED ISOPERIMETRIC PROBLEMS - - - - -	19
By OSKAR BOLZA, Professor of Mathematics	
IV. TERNARY ORTHOGONAL GROUP IN A GENERAL FIELD - - - - -	27
By LEONARD EUGENE DICKSON, Assistant Professor of Mathematics	
V. THE GROUPS DEFINED FOR A GENERAL FIELD BY THE ROTATION GROUPS - - - - -	35
By LEONARD EUGENE DICKSON, Assistant Professor of Mathematics	
VI. ON AMORPHOUS SULPHUR - - - - -	53
By ALEXANDER SMITH, Associate Professor of General Chemistry, with the collaboration of WILLIS B. HOLMES	
VII. NEW INSTRUMENTS OF PRECISION FROM THE RYERSON LABORATORY - - - - -	65
By ROBERT A. MILLIKAN, Assistant Professor of Physics	
VIII. ON THE "BECKMANN REARRANGEMENT" - - - - -	73
By JULIUS STIEGLITZ, Associate Professor of Chemistry	
IX. A SYSTEMATIC METHOD OF CALCULATING THE DIMENSIONS OF DIRECT-CURRENT DYNAMO-ELECTRIC MACHINES - - - - -	89
By CARL KINSLEY, Instructor in Physics	

- X. EQUILIBRIUM IN THE SYSTEM COMPOSED OF SODIUM CARBONATE, SODIUM  
BICARBONATE, CARBON DIOXIDE, AND WATER - - - 103  
By HERBERT N. MCCOY, Instructor in Chemistry
- XI. INVARIANTS AND COVARIANTS OF QUADRATIC DIFFERENTIAL QUANTICS  
OF  $n$  VARIABLES - - - 125  
By HEINRICH MASCHKE, Associate Professor of Mathematics
- XII. THE SUBGROUPS OF THE GENERALIZED FINITE MODULAR GROUP - 139  
By ELIAKIM HASTINGS MOORE, Professor and Head of the Department  
of Mathematics
- XIII. A CONTRIBUTION TO THE THEORY OF GLACIAL MOTION - - - 191  
By THOMAS CHROWDER CHAMBERLIN, Professor and Head of the Depart-  
ment of Geology

## **THE VELOCITY OF LIGHT**





# THE VELOCITY OF LIGHT

ALBERT A. MICHELSON

THE fact that the velocity of light is so far beyond the conception of the human intellect, coupled with the extraordinary accuracy with which it may be measured, makes this determination one of the most fascinating problems that fall to the lot of the investigator.

The experiment of measuring the velocity of light was originally devised by Arago, as a means of furnishing a crucial test between the two great rival theories of light. History has recorded the triumph of the undulatory theory, as the result of the classic experiments of Foucault and Fizeau.

It was found in the course of these and subsequent experiments that the experimental determination of the absolute value ( $V$ ) of the velocity of light was capable of so high a degree of accuracy that, instead of obtaining  $V$  by observations of the eclipses of Jupiter's satellites, or the aberration of the fixed stars, together with the astronomically calculated value of the Sun's distance (with results of relatively small accuracy), the process is exactly reversed; so that the experimental value of  $V$ , together with observed values of the light equation, or the constant of aberration, furnishes a far more accurate value of the Sun's distance than can be obtained by rare and costly expeditions for observing the transit of Venus. This appears quite clearly if we contrast the numbers given by Professor Harkness,<sup>1</sup> which give the solar parallax as obtained from astronomical data,<sup>2</sup>  $8.78 \pm .05$ , with the results given by Professor Todd of the same quantity as determined by the combination of velocity of light with the light equation or with the constant of aberration.

In an article on "Solar Parallax from the Velocity of Light,"<sup>3</sup> Professor D. P. Todd gives the following résumé of the determinations of the solar parallax from  $L$ , the light equation, and from  $\alpha$ , the constant of aberration:

The elements of sensible uncertainty considered are:

1. Uncertainty in the determination of terrestrial velocity of light. . . . I am disposed to think that the limit of uncertainty of the velocity of light concluded above (299,920 kilometers) may be fairly taken at seventy kilometers.
2. Uncertainty in the coefficient of the light equation from observations of the satellites of Jupiter. . . . The amount of uncertainty is probably not far from one second of time.<sup>4</sup>
3. Uncertainty in the constant of sidereal aberration. I conceive that a variation of 0.025 in this well-determined constant<sup>5</sup> will not be regarded far from the limit of uncertainty. . . .
4. Uncertainty in the relation of the absolute terrestrial velocity to the velocity in space.

<sup>1</sup>The numbers here given are not Professor Harkness's estimates, but only a rather rough guess from the data he presents, which I think are nevertheless of about the right order of magnitude.

<sup>2</sup>*Amer. Jour. Sci.*, XXII (1881), p. 375.

<sup>3</sup>*Ibid.*, XIX (1880), p. 59.

<sup>4</sup>This is taken from the observations of Delambre, 493.2, weight 1, and Glasenapp, 500.84  $\pm$  1.02, weight 2.

<sup>5</sup>Professor Todd accepts Struve's value, 2'0445.

. . . . The impossibility of an experimental determination of this relation renders the assumption of identity necessary.

In conclusion, then, all the experimental determinations of the velocity of light hitherto made give, when combined with astronomical constants, the mean equatorial horizontal parallax of the Sun,  $8'808 \pm 0.006$ . The corresponding mean radius of the terrestrial orbit is 149,345,000 kilometers.

Roughly speaking, it appears that the velocity-of-light method of obtaining the solar parallax is about ten times as accurate as the astronomical method. It also appears that in the last-named method the order of accuracy in the determination of the astronomical factors is about one one-thousandth part.

The undulatory theory does not specify the nature of the undulations, and it is no argument against it that the special mode of motion assumed by Fresnel (vibrations of an elastic solid, which so elegantly explains all the known phenomena of light) has met with some serious objections. If, according to Maxwell, these undulations are electromagnetic changes, these objections no longer apply, and this electromagnetic theory of light is now universally accepted.\*

In accordance with Maxwell's theory, the ratio of any electric or magnetic unit measured electrostatically to the same unit measured electromagnetically, the ratio usually designated as Maxwell's  $v$ , must be equal to the velocity of light. The verification of this relation would be a powerful argument (were any such needed) in favor of the theory; and, in fact, the value of  $v$  agrees as well with  $V$  as do the separate accepted values of  $v$  among themselves.

M. Abraham, in a report to the International Congress of Physics, gives the following résumé of these determinations:

Himstedt	-	-	-	-	-	-	-	$3.0057 \times 10^{10}$
Rosa	-	-	-	-	-	-	-	$3.0000 \times 10^{10}$
J. J. Thomson	-	-	-	-	-	-	-	$2.9960 \times 10^{10}$
H. Abraham	-	-	-	-	-	-	-	$2.9913 \times 10^{10}$
Pellat	-	-	-	-	-	-	-	$3.0092 \times 10^{10}$
Hurmuzescu	-	-	-	-	-	-	-	$3.0010 \times 10^{10}$
Perot and Fabry	-	-	-	-	-	-	-	$2.9973 \times 10^{10}$

M. Abraham accepts the mean value  $v = 3.0001 \times 10^{10}$ , which he regards as probably correct to the thousandth part, and concludes with the following significant remark:

Étant donné l'intérêt qui s'attache à la détermination de la vitesse  $v$ , il paraît désirable que de nouvelles expériences soient entreprises. La précision des anciennes mesures peut être dépassée: toutes les méthodes s'y prêtent. Il y a encore à réduire quelques corrections trop incertaines; il y a à simplifier quelques mesures auxiliaires trop complexes, et par ce nouvel effort on pourra, sans aucune doute, apporter dans la mesure de  $v$  une précision supérieure à celle aujourd'hui acquise pour la vitesse de la lumière.

\* While the electromagnetic theory gives a satisfactory explanation of light, and goes a long way toward the explanation of the mechanism of radiation, yet it seems not at

all unlikely that a return to something analogous to the now discarded electro-solid theory may be necessary to explain electromagnetic phenomena.

Again, Blondlot and Gutton, in the same report, give the following résumé of some of the best determinations of the velocity of Hertzian waves :

Blondlot	- - - - -	302,200 km./sec.
Blondlot	- - - - -	{ 296,400
		{ 298,000
Trowbridge and Duane	- - - - -	300,300
MacLean	- - - - -	299,110
Saunders	- - - - -	{ 298,200
		{ 299,700

Messrs. Blondlot and Gutton do not make any estimate of the most probable value, but doubtless the mean of the preceding results,  $299,130 \pm 1,000$ , cannot be far from the truth. The authors conclude as follows :

Ils [ces valeurs] sont sensiblement égaux au rapport des unités électromagnétiques et électrostatiques de quantité d'électricité, comme la théorie de Maxwell l'indique, et aussi à la vitesse de la lumière. Il y a un intérêt capital à rechercher si cette égalité des trois nombres est seulement approximative, ou si elle offre un caractère absolu ; on devra donc s'efforcer, dans les déterminations ultérieures de la vitesse de propagation des ondes électromagnétiques, non seulement d'atteindre la plus grande exactitude possible, mais aussi de déterminer le degré d'approximation des nombres obtenus.

It would appear, then, that there is but little doubt that in the near future both these determinations will be made with at least the same high order of accuracy as obtains in the measurement of the velocity of light. In this case any possible difference in the resulting values would not cast any doubt upon the electromagnetic theory, but would doubtless be traceable to the enormous difference in the conditions determined by light-waves on the one hand and electric oscillations or static charges on the other. On the contrary, such a difference might almost certainly be predicted, and would probably throw much light on the structure and mode of action of dielectrics.

Having reviewed in some detail the order of accuracy attained in the astronomical elements which are to be combined with  $V$ , and the electromagnetic results which are to be compared with  $V$ , let us consider the actual order of accuracy obtained in the measurement of  $V$  itself.

Concerning this, as in the cases just reviewed, a considerable diversity of opinion will prevail concerning the relative weight which should be attributed to the various determinations. The following table is taken from Professor Newcomb's report :

1. Foucault, 1862	- - - - -	298,000
2. Cornu (1), 1874	- - - - -	298,500
3. Cornu (2), 1878	- - - - -	300,400
4. Cornu (2), as discussed by Listing	- - - - -	299,990
5. Young and Forbes, 1880-81	- - - - -	301,382
6. Michelson (1)	- - - - -	299,910
7. Michelson (2)	- - - - -	299,853
8. Newcomb (selected results)	- - - - -	299,860
9. Newcomb (all observations)	- - - - -	299,810

No doubt the weight which should be assigned to these results varies enormously. Still, in view of the limited number, and keeping in mind the criticism of M. Cornu, which will be presently referred to, it may be of interest to find the mean of these nine values, giving them equal weight. The result is  $V = 299,664$ , with an average difference from the mean of 600 kilometers.

But the results 1, 2, 6 are admitted to be only first approximations, and 5 is undoubtedly affected by some serious constant error; so that a much closer approximation to the true value will be obtained by combining 3 or 4, the mean of 6 and 7, and 9. The first combination gives:

Cornu	-	-	-	-	-	-	-	-	300,400 $\pm$ 300
Michelson	-	-	-	-	-	-	-	-	299,882 $\pm$ 60
Newcomb	-	-	-	-	-	-	-	-	299,810 $\pm$ 60

Giving these equal weight, we find  $V = 300,030$ , with an average difference from the mean of about 250.

If, however, we take Listing's valuation of Cornu's results, we get:

Cornu	-	-	-	-	-	-	-	-	299,990
Michelson	-	-	-	-	-	-	-	-	299,882
Newcomb	-	-	-	-	-	-	-	-	299,810

Giving for the mean  $V = 299,890$ , with an average difference from the mean of about 60.

In a paper presented to the International Congress in 1900, M. Cornu expresses serious doubts concerning the order of accuracy of the results obtained by means of the revolving mirror, and, giving equal weight to the two methods regardless of the number of determinations, and ignoring Listing's discussion of his own results, he gives, as his estimate of the most probable value of the velocity of light,  $V = 300,130$ , with an average difference from the mean of 270.

The criticism of M. Cornu embraces the following points:

1. Can the ordinary laws of reflection be applied to the case of a mirror in rapid rotation?
2. Can the ordinary laws of reflection be applied to the case of a light-ray whose axis moves with a velocity comparable with the velocity of light?
3. Does the dragging along (*entrainement*) of the light-waves in the air vortex close to the mirror affect the displacement?

These criticisms have been carefully considered by H. A. Lorentz,<sup>1</sup> who states that, "Quoiqu'elles [quelques considérations . . .] laissent encore beaucoup à désirer au point de vue de la rigueur, elles me semblent bien propres à dissiper les doutes que je viens de rappeler."

It seems to me that M. Lorentz has satisfactorily answered M. Cornu's questions; and my own estimate of the most probable value of  $V$  agrees with the second result

<sup>1</sup> "Sur la méthode du miroir tournant pour la détermination de la vitesse de la lumière," *Arch. Néer des Sciences, Sér. II*, T. VI, p. 303.

given above, namely,  $V = 299,890 \pm 60$ . Nevertheless, an experimental determination which should more or less completely avoid the necessity for the consideration of these questions, or at least materially diminish any error which could possibly arise, even if they were valid, and which, in any case, would furnish one more independent result to be combined with the very few reliable ones already obtained, would surely be welcome.

The preceding exposition makes it clear that the measurements and observations to be combined with, or to be compared with, the velocity of light—namely, (1) astronomical observations of  $L$ ; (2) electromagnetic ratio  $v$ ; (3) velocity of Hertzian oscillations—are all capable of measurement with an order of accuracy which even now approaches, if it does not equal, that of the measurement of the velocity of light. This, it seems to me, is a sufficient excuse—if any were needed—for once more attacking the problem.

The following plan suggested itself during the experiments upon the “relative motion of the earth and the luminiferous ether”:<sup>8</sup>

The essential feature is the combination of a grating with a revolving mirror, which combination acts as a toothed wheel; the grating space representing the distance between the teeth, the radius being the distance from the revolving mirror to the grating.

It was proposed to utilize this combination in an attempt to solve the problem of the “relative motion” by measuring the velocity of light *in one direction*, that is, without returning the light to the source. This was before the celebrated work of Hertz showed that the electrical impulses (which were to be used to establish the required phase relation between the two revolving mirrors) would be affected in the same way as would the light-waves themselves. It may also be noted in this connection that the method proposed in the same article, and illustrated by Fig. 1, is also not sound.<sup>9</sup>

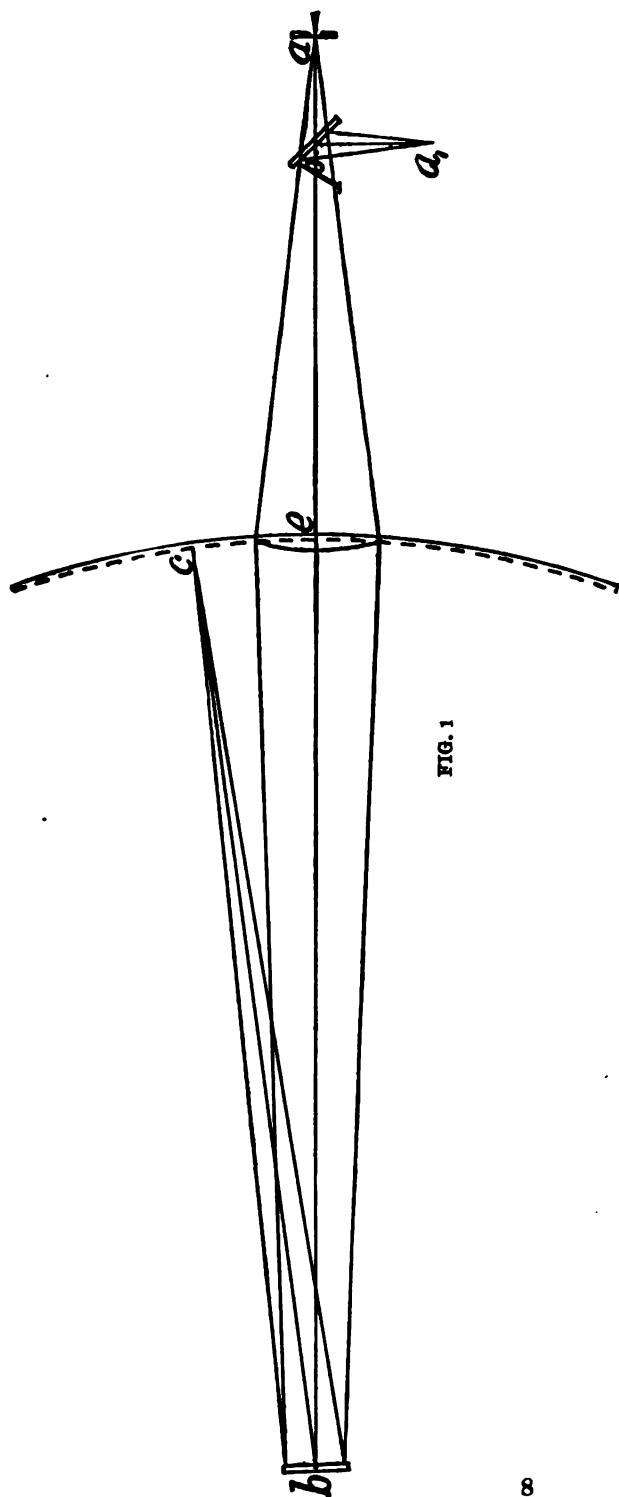
The plan proposed is virtually a combination of the methods of Foucault and Fizeau; the essential feature of the observation of eclipses corresponding to the latter method, while the production of the eclipses is brought about by a revolving beam of light as in the former method.

Figs. 1 and 2, in which the lettering is the same, will illustrate the essential features, subject to such minor modifications as experiment may suggest. The light starts from a slit at  $a$ , passes through a lightly silvered glass plate  $P$  and a lens  $e$ , and falls upon the upper half of the revolving mirror  $b$ . Thence it proceeds to the grating  $c$ , upon the surface of which it forms an image of the slit. It is thence reflected to the lower half  $d$  of the revolving mirror, which reflects the beam through the lens  $f$  to the distant mirror upon the surface of which the second image of the slit is formed.

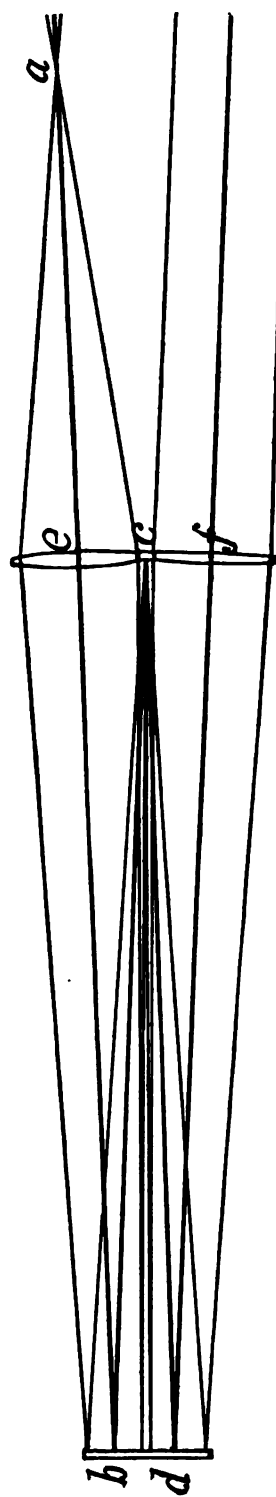
<sup>8</sup> *Amer. Jour. Sci.*, XXXIV (1887), November.

<sup>9</sup> It may also be worth mentioning that some preliminary experiments, made about two years ago, have shown that it is not entirely impossible to employ a mechanical method of keeping the two revolving mirrors in a con-

stant phase relation. For instance, it was shown that the vibrations of a tuning-fork could be transmitted over a mile of piano wire with a diminution of amplitude of less than one-half.



8



The light then retraces its course and returns to its source at  $a$ , part being reflected to  $a_1$  for convenience of observation by the eyepiece. The limit of closeness of the grating space is determined by the aperture of the revolving mirror viewed from  $c$ .

If the "radius"  $bc$  is 3 m. and the revolving mirror is 6 cm. wide, this angle will be 0.02; and the breadth of the diffraction image at  $c$  will be of the order  $\lambda/.02$ , or, say, 0.025 mm. The grating space should therefore be at least 0.05 mm., and, probably better, 0.1 mm. If the number of revolutions is 250, and the distance to the fixed mirror be 3 km., the displacement of the first image over the grating surface will be 18 cm., corresponding to 1,800 eclipses. There need be no difficulty in counting the order of the eclipse observed, if the speed is gradually increased to its final value. The fractions could probably be observed correctly to something like 2 per cent., so that this element of the computation for  $V$  could be measured to something like 1 part in 100,000.

This same, or even a higher, order of accuracy may be obtained in the measurement of all the other elements. Previous experiments have shown that the speed of the mirror may be obtained by means of a rated tuning-fork to within 1 in 100,000. The measurement of the distance may be made directly on a base line especially prepared for such work to within 1 in 200,000.

Finally, the grating may be calibrated to an order of accuracy depending on the angle subtended by the diffraction fringes, *i. e.*,  $\frac{1}{40} \times \frac{1}{30000}$ , or less than 1 in 100,000.

It seems not unreasonable to hope that, with proper care and patience, the value of this great fundamental constant of nature may be found to within 5 km. or less.

It will be noted that in the form of experiment here proposed the more serious of the difficulties pointed out by M. Cornu no longer exist, or are very much diminished, while the possible accuracy is greatly increased. It may be hoped, therefore, that the result of this combination of the methods of Foucault and Fizeau will be to reconcile the differences which thus far seem to exist between the results of the work of their respective followers.





**THE ISOPERIMETRIC PROBLEM ON A  
GIVEN SURFACE**



# CONCERNING THE GEODESIC CURVATURE AND THE ISOPERIMETRIC PROBLEM ON A GIVEN SURFACE

OSKAR BOLZA

IN the following note I propose to give a solution of the isoperimetric problem on a surface—as far as the first and second necessary conditions are concerned—which seems to be considerably simpler, and at the same time more general, than the solutions recently published by Kneser<sup>1</sup> and Whittemore.<sup>1</sup> For this purpose I use, instead of Bonnet's expression for the geodesic curvature of a curve traced on a surface, a formula expressing the geodesic curvature when the curve is given in parameter representation.

As this formula is not to be found in the usual treatises<sup>2</sup> on differential geometry, I give in § 1 an apparently new and elementary proof of it, paying special attention to the question of sign, which is essential for the discussion of the second variation.

Then I proceed, in § 2, to the consideration of the isoperimetric problem on a surface, with fixed endpoints, including a discussion of the second necessary condition for a maximum (Legendre's criterion); and finally I consider the case of variable endpoints.

## §1. THE GEODESIC CURVATURE

The *geodesic curvature* of a curve  $C$  traced on a surface  $S$  in a point  $P_0$  of  $C$  is defined as the curvature, in  $P_0$ , of the projection  $C'$  of  $C$  upon the tangent plane in  $P_0$  to the surface  $S$ ; the corresponding center of curvature  $M_0$  of  $C'$  is called the *center of geodesic curvature*.<sup>3</sup>

Let  $x = \phi(u, v)$ ,  $y = \psi(u, v)$ ,  $z = \chi(u, v)$  (1)

represent the surface  $S$ , referred to a system of rectangular co-ordinates, with the customary agreements<sup>4</sup> concerning the positive directions of the axes; let, further,

$u = u(t)$ ,  $v = v(t)$  (2)

represent the curve  $C$ , and let  $t = t_0$  be the parameter of  $P_0$ ;  $x_0, y_0, z_0$  its co-ordinates.

Then, in order to obtain the expression for the geodesic curvature, we introduce a new rectangular system of co-ordinates,  $\xi, \eta, \zeta$ , whose origin is the point  $P_0$ , whose

<sup>1</sup>KNESER, *Lehrbuch der Variationsrechnung*, § 34; WHITTEMORE, "The Isoperimetrical Problem on Any Surface," *Annals of Mathematics* (2), Vol. II, p. 176. Kneser supposes the given curve, Whittemore also the unknown curve, in the form  $v = f(u)$ , which is an essential restriction. Both use a normal form for the line element, which implies another, though less important, restriction. By entirely different methods the problem is treated by DARBOUX, *Théorie des surfaces*, Vol. III, Nos. 650, 651.

<sup>2</sup>It is, however, given in LAURENT, *Traité d'Analyse*, Vol. VII, p. 132, and in an article by MINDING, *Journal für Mathe-*

*matik*, Vol. LXXXVI, p. 281, in both places without a discussion of the sign; a paper by STAUBE concerning the sign of the geodesic curvature (*Sitzungsberichte der Naturforschergesellschaft bei der Universität Dorpat*, Vol. XI, p. 72) has not been accessible to me.

<sup>3</sup>DARBOUX, *loc. cit.*, Vol. II, No. 490; BIANCHI, *Vorlesungen über Differentialgeometrie*, § 75.

<sup>4</sup>A left-hand rotation through 90° around the positive  $x$ -axis brings the positive  $x$ -axis to coincidence with the positive  $y$ -axis.

positive  $\xi$ -axis coincides with the positive direction of the tangent to  $C$  in  $P_0$ , whose positive  $\zeta$ -axis coincides with the positive<sup>5</sup> direction of the normal to the surface in  $P_0$ , and whose positive  $\eta$ -axis lies to the left of the positive  $\xi$ -axis, if we look down upon the  $\xi, \eta$ -plane from the positive  $\zeta$ -axis.

Let  $x, y, z$  and  $\xi, \eta, \zeta$  be the co-ordinates of the same point  $P(t)$  of  $C$  in the old and the new systems, respectively; then

$$\begin{aligned}\xi &= l(x - x_0) + l_1(y - y_0) + l_2(z - z_0) \\ \eta &= m(x - x_0) + m_1(y - y_0) + m_2(z - z_0) \\ \zeta &= n(x - x_0) + n_1(y - y_0) + n_2(z - z_0),\end{aligned}\quad (3)$$

where the coefficients are the well-known direction-cosines.

The two first equations, together with  $\zeta = 0$ , will then represent the curve  $C'$  in terms of the parameter  $t$ , and therefore the geodesic curvature of  $C$  at  $P_0$  will be given by the expression<sup>6</sup>

$$\frac{1}{\rho_g} = \frac{\xi' \eta'' - \xi'' \eta'}{(\sqrt{\xi'^2 + \eta'^2})^3}, \quad \text{for } t = t_0, \quad (4)$$

where the accents indicate derivation with respect to  $t$ , and the square root is understood positive, as all the square roots in the sequel. The sense of the vector  $P_0 M_0$  is indicated by the sign of  $\frac{1}{\rho_g}$ : if it is  $+$ ,  $P_0 M_0$  lies to the left of the positive tangent to  $C'$  at  $P_0$ ; if  $-$ , to the right.

But

$$\xi' \eta'' - \xi'' \eta' = \begin{vmatrix} l & l_1 & l_2 \\ m & m_1 & m_2 \end{vmatrix} \begin{vmatrix} x' & y' & z' \\ x'' & y'' & z'' \end{vmatrix},$$

or, since

$$l_1 m_2 - l_2 m_1 = n = \frac{y_u z_v - y_v z_u}{\sqrt{EG - F^2}}, \text{ etc.,}$$

where the subscripts  $u, v$  denote, as always in the sequel, partial derivation with respect to  $u$  and  $v$ , respectively:

$$(\xi' \eta'' - \xi'' \eta') \sqrt{EG - F^2} = \begin{vmatrix} x_u & y_u & z_u \\ x_v & y_v & z_v \end{vmatrix} \begin{vmatrix} x' & y' & z' \\ x'' & y'' & z'' \end{vmatrix}.$$

But

$$\begin{aligned}x' &= x_u u' + x_v v' \\ x'' &= x_u u'' + x_v v'' + x_{uu} u'^2 + 2x_{uv} u' v' + x_{vv} v'^2,\end{aligned}$$

and if we denote, as usual,

$$\sum x_u^2 = E, \quad \sum x_u x_v = F, \quad \sum x_v^2 = G,$$

the summation referring to a cyclic permutation of  $x, y, z$ , then

$$\begin{aligned}\sum x_u x_{uu} &= \frac{1}{2} E_u, \quad \sum x_u x_{uv} = \frac{1}{2} E_v, \quad \sum x_v x_{uv} = \frac{1}{2} G_u, \quad \sum x_v x_{vv} = \frac{1}{2} G_v, \\ \sum x_u x_{v^2} &= F_v - \frac{1}{2} G_u, \quad \sum x_v x_{u^2} = F_u - \frac{1}{2} E_v.\end{aligned}$$

<sup>5</sup> BLANCHI, *loc. cit.*, §§ 34, 46.

<sup>6</sup> See, for instance, C. JORDAN, *Cours d'Analyse*, Vol. I, No. 443.

Substituting these values, we obtain

$$\xi' \eta'' - \xi'' \eta' = \frac{\Gamma}{\sqrt{EG - F^2}},$$

where

$$\begin{aligned} \Gamma = & (EG - F^2)(u'v'' - u''v') + \\ & (Eu' + Fv')\left(u'^2(F_u - \frac{1}{2}E_v) + u'v'G_u + \frac{1}{2}v'^2G_v\right) - \\ & (Fu' + Gv')\left(\frac{1}{2}u'^2E_u + u'v'E_v + v'^2(F_v - \frac{1}{2}G_u)\right). \end{aligned} \quad (5)$$

On the other hand, we have, for  $t = t_0$ ,  $\xi' = 0$ , and therefore

$$\xi'^2 + \eta'^2 = \xi''^2 + \eta''^2 + \zeta'^2 = x'^2 + y'^2 + z'^2 = Eu'^2 + 2Fu'v' + Gv'^2.$$

Hence we have reached the result:

*The geodesic curvature of the curve C in  $P_0$  has the value*

$$\frac{1}{\rho_g} = \frac{\Gamma}{\sqrt{EG - F^2}(\sqrt{Eu'^2 + 2Fu'v' + Gv'^2})^3}, \quad (6)$$

where  $\Gamma$  is defined by (5).

The sign of  $\Gamma$  decides the *sense* of the vector from  $P_0$  to the center of geodesic curvature  $M_0$ .

*Corollary I.*—If  $\Gamma > 0$  the vector  $P_0M_0$  is to the left of the positive tangent to  $C$ , when we look down upon the tangent plane from the positive normal to  $S$  in  $P_0$ ; if  $\Gamma < 0$ , to the right.

*Corollary II.*—The equation

$$\Gamma = 0$$

is the *differential equation of the geodesic lines.*<sup>1</sup>

It is obtained in this form when we define the geodesics by their minimum property. For, if we denote by  $\Phi(f)$  the operator

$$\Phi(f) = f_{uv} - f_{vu} + f_1(u'v'' - u''v'), \quad (7)$$

where  $f$  is a positively<sup>2</sup> homogeneous function of  $u, v, u', v'$  of dimension 1, and

$$f_1 = \frac{f_{uu'}}{v'^2} = -\frac{f_{uv'}}{u'v'} = \frac{f_{v'v'}}{u'^2},$$

the first necessary condition for a minimum of the integral

$$I = \int_{t_0}^{t_1} \sqrt{Eu'^2 + 2Fu'v' + Gv'^2} dt$$

is, according to Weierstrass,<sup>3</sup>

$$\Phi(\sqrt{Eu'^2 + 2Fu'v' + Gv'^2}) = 0.$$

<sup>1</sup> KNOBLAUCH, *loc. cit.*, § 53(9); BIANCHI, *loc. cit.*, § 78(10).

<sup>2</sup> *Lectures on the Calculus of Variations.*

<sup>3</sup> See KNEBES, *loc. cit.*, § 2.

Since here

$$f_1 = \frac{EG - F^2}{(\sqrt{Eu'^2 + 2Fu'v' + Gv'^2})^3},$$

this reduces to

$$\Phi(\sqrt{Eu'^2 + 2Fu'v' + Gv'^2}) = \frac{\Gamma}{(\sqrt{Eu'^2 + 2Fu'v' + Gv'^2})^3},$$

and therefore the geodesic curvature may also be written :

$$\frac{1}{\rho_g} = \frac{\Phi(\sqrt{Eu'^2 + 2Fu'v' + Gv'^2})}{\sqrt{EG - F^2}}. \quad (8)$$

## §2. THE ISOPERIMETRIC PROBLEM ON A GIVEN SURFACE

Suppose that on the surface of  $S$  there is given a curve  $\tilde{C}$  :

$$u = \tilde{u}(\tau), \quad v = \tilde{v}(\tau),$$

and on it two points  $A(\tau_0)$  and  $B(\tau_1)$ ; it is required to join the points  $A$  and  $B$  by a curve  $C$  on  $S$  :

$$u = u(t), \quad v = v(t)$$

of given length  $l$  such that the area of the region  $R$  bounded by  $C$  and  $\tilde{C}$  shall be a maximum.

Denoting the parameters of  $A$  and  $B$  on  $C$  by  $t_0$  and  $t_1$ , respectively, where we may always suppose  $t_0 < t_1$ , and by  $N$  a function of  $u$  and  $v$  whose partial derivative with respect to  $v$  is  $\sqrt{EG - F^2}$  :

$$N_v = \sqrt{EG - F^2}, \quad (9)$$

we have to maximize the integral <sup>10</sup>

$$I = \epsilon \int_{t_0}^{t_1} N(u, v) u' dt + \epsilon \int_{\tau_1}^{\tau_0} N(\tilde{u}, \tilde{v}) \tilde{u}' d\tau, \quad (10)$$

where  $\epsilon = \pm 1$ , with the condition that the integral

$$K = \int_{t_0}^{t_1} \sqrt{Eu'^2 + 2Fu'v' + Gv'^2} dt \quad (11)$$

shall have the given value  $l$ .

The sign of  $\epsilon$  has to be determined as follows: Let  $R'$  be the image of the region  $R$  in the  $u, v$ -plane, whose positive  $v$ -axis is supposed to be to the left of the positive  $u$ -axis; let the point  $u, v$  describe the boundary of  $R'$  counter-clockwise, starting from the image  $A'$  of  $A$ ; then, as the corresponding point  $x, y, z$  describes the boundary of  $R$ , starting from  $A$ , it will either go first along  $C$  or first along  $\tilde{C}$ . In the first case  $\epsilon = -1$ , in the second  $\epsilon = +1$ . This follows from Green's Theorem, which has been applied in establishing the expression for  $I$ .

<sup>10</sup> For a detailed statement of the conditions which must be imposed upon the surface  $S$  and the curves  $C$  and  $\tilde{C}$ , in order that the area in question may exist and be

expressible by (10), I refer to C. JORDAN, *Cours d'Analyse*, Vol. I, Nos. 102, 112, 155-7, and Vol. II, Nos. 132, 133.

Since the second integral in (10) is constant, the first necessary condition for a maximum is, in Weierstrass's form and in the notation of § 1 (7),

$$\Phi(H) = 0, \quad (12)$$

$$\text{where} \quad H = \epsilon N u' + \lambda \sqrt{E u'^2 + 2F u' v' + G v'^2}, \quad (18)$$

$\lambda$  being the isoperimetric constant.

But

$$\Phi(H) = \epsilon \Phi(N u') + \lambda \Phi(\sqrt{E u'^2 + 2F u' v' + G v'^2}),$$

and

$$\begin{aligned} \epsilon \Phi(N u') &= -\epsilon N_v = -\epsilon \sqrt{E G - F^2} \\ \Phi(\sqrt{E u'^2 + 2F u' v' + G v'^2}) &= \frac{\sqrt{E G - F^2}}{\rho_g}, \end{aligned}$$

according to (8).

Hence the differential equation (18) reduces to:

$$\rho_g = \epsilon \lambda; \quad (14)$$

that is: *The extremals are curves of constant geodesic curvature (geodesic circles).*<sup>11</sup>

The direction of the vector from a point  $P$  of  $C$  to the center of geodesic curvature  $M$  is determined by Legendre's condition, according to which, for a maximum, the function

$$H_1 = \frac{H_{u'u'}}{v'^2}$$

must not be positive on  $C$ .

But

$$H_1 = \frac{\lambda (E G - F^2)}{(\sqrt{E u'^2 + 2F u' v' + G v'^2})^3}.$$

Hence: For a maximum  $\lambda$  must be negative, and therefore  $\rho_g$  must have the opposite sign of  $\epsilon$ .

The direction of the vector  $PM$  is then unambiguously determined by the rule of sign given in § 1.

If only one of the endpoints, say  $A$ , is given, while the other is to be determined on  $\tilde{C}$ , all the previous results hold for variations which leave  $B$  unchanged, and therefore the extremals are geodesic circles. For a variation which changes  $B$  into a point  $\bar{B}$  on  $\tilde{C}$ , with the parameter  $\tau_1 + h$ , the total variation of the area is

$$\Delta I = \Delta \int_{t_0}^{t_1} \epsilon N(u, v) u' dt + \int_{\tau_1+h}^{\tau_1} \epsilon N(\bar{u}, \bar{v}) \bar{u}' d\tau.$$

Apply the usual process<sup>12</sup> to the first integral, and then expand the whole expression according to powers of  $h$ ; the coefficient of  $h$  must vanish, i. e.,

$$H_{u'} \bar{u}' + H_v \bar{v}' - \epsilon \tilde{N} \bar{u}' = 0,$$

where  $\tilde{N}$  stands for  $N(\bar{u}, \bar{v})$ .

<sup>11</sup> See DARBOUX, *loc. cit.*, Vol III, p. 151.

<sup>12</sup> KNEBEL, *loc. cit.*, § 33.



And, since  $u(t_1) = \bar{u}(\tau_1)$ ,  $v(t_1) = \bar{v}(\tau_1)$ ,  $N$  cancels out, and the condition reduces to

$$(Eu' + Fv')\bar{u}' + (Fu' + Gv')\bar{v}' = 0,$$

which expresses that  $C$  must be orthogonal to  $\bar{C}$  in  $B$ .

The following modification of the problem deserves to be mentioned: Suppose not the length of  $C$  is given, but the length of the whole contour  $C$  and  $\bar{C}$ . In this case

$$K = \int_{t_0}^{t_1} \sqrt{Eu'^2 + 2Fu'v' + Gv'^2} dt + \int_{\tau_1}^{\tau_0} \sqrt{E\bar{u}'^2 + 2F\bar{u}'\bar{v}' + G\bar{v}'^2} d\tau,$$

if  $\tau_1 < \tau_0$ . If  $A$  and  $B$  are both given, the second integral in  $K$  is constant, and the analytic problem is the same as before; but if  $B$  is movable on  $\bar{C}$ , we obtain:

$$\Delta K = \Delta \int_{t_0}^{t_1} \sqrt{Eu'^2 + 2Fu'v' + Gv'^2} dt + \int_{\tau_1+h}^{\tau_1} \sqrt{E\bar{u}'^2 + 2F\bar{u}'\bar{v}' + G\bar{v}'^2} d\tau.$$

The process indicated above now leads to the result:

$$H_u \bar{u}' + H_v \bar{v}' - \bar{H} \bar{u}' = 0,$$

or

$$\frac{(Eu' + Fv')\bar{u}' + (Fu' + Gv')\bar{v}'}{\sqrt{Eu'^2 + 2Fu'v' + Gv'^2} \sqrt{E\bar{u}'^2 + 2F\bar{u}'\bar{v}' + G\bar{v}'^2}} = 1,$$

whose geometrical meaning is that, in this case,  $C$  must be tangent to  $\bar{C}$  in  $B$ .

**THE SECOND VARIATION IN THE SO-CALLED  
ISOPERIMETRIC PROBLEMS**



# PROOF OF THE SUFFICIENCY OF JACOBI'S CONDITION FOR A PERMANENT SIGN OF THE SECOND VARIATION IN THE SO-CALLED ISOPERIMETRIC PROBLEMS

OSKAR BOLZA

## §1. FORMULATION OF THE PROBLEM

THE discussion of the second variation for the simplest class of isoperimetric problems,<sup>1</sup> in parameter representation, leads to the following question:

Let  $H_1, H_2, T$  be three given functions<sup>2</sup> of  $t$ , regular in an interval<sup>3</sup>  $(t_0, t_1)$ ; moreover, it is supposed that  $H_1 > 0$  and  $T \neq 0$  on  $(t_0, t_1)$ . Under what conditions will the definite integral

$$\delta^2 I = \int_{t_0}^{t_1} \left[ H_1 \left( \frac{dw}{dt} \right)^2 + H_2 w^2 \right] dt$$

be positive for all functions  $w$ , not identically zero, satisfying the following conditions:

(a)  $w(t_0) = 0, \quad w(t_1) = 0,$

(b)  $\int_{t_0}^{t_1} w T dt = 0,$

(c) the functions  $w$  satisfy certain conditions concerning continuity and existence and continuity of the first derivative.

With respect to (c) we make the assumption<sup>4</sup> that  $w$  itself shall be continuous on  $(t_0, t_1)$ , and that the interval  $(t_0, t_1)$  can be divided into a finite number of subintervals such that on each subinterval the first derivative exists and is continuous—with the understanding that in the lower (upper) endpoint of each subinterval “progressive (regressive) derivative” is substituted for “derivative.”

The answer is as follows: Denote by  $\Psi(w)$  the differential expression

$$\Psi(w) = H_2 w - \frac{d}{dt} \left( H_1 \frac{dw}{dt} \right),$$

<sup>1</sup>The class treated in KNEBESER's *Lehrbuch der Variationsrechnung* in chap. iv, in which it is required to minimize an integral of the form

$$I = \int_{t_0}^{t_1} F(x, y, x', y') dt,$$

while at the same time another integral of the same form

$$K = \int_{t_0}^{t_1} G(x, y, x', y') dt$$

has a prescribed value.

<sup>2</sup> $H_1, H_2$  are derived from  $H = F + \lambda G$  in the same man-

ner as  $F_1, F_2$  from  $F$  in the theory of the unconditioned problem; for Weierstrass's explicit expression of  $F_2$  see BLISS, *Transactions of the American Mathematical Society*, Vol. III (1902), p. 133. Further

$$T = G_{xy'} - G_{x'y} + G_1(x'y' - x'y'),$$

the literal subscripts denoting partial differentiation.

<sup>3</sup>The notation implies that  $t_0 < t_1$ .

<sup>4</sup>This agreement means for the isoperimetric problem that we restrict ourselves to continuous curves made up of a finite number of arcs, along each of which the curve has a continuously turning tangent.

#### 4 THE SECOND VARIATION IN THE SO-CALLED ISOPERIMETRIC PROBLEMS

and by  $U$  and  $V$  two integrals<sup>5</sup> of the differential equations:

$$\Psi(U) = 0, \quad \Psi(V) = T, \quad (1)$$

respectively, which vanish for  $t = t_0$ :

$$U(t_0) = 0, \quad V(t_0) = 0; \quad (2)$$

further let

$$M = \int_{t_0}^{t_1} U T dt, \quad N = \int_{t_0}^{t_1} V T dt,$$

and put

$$\Delta(t, t_0) = M V - N U.$$

Then, if  $t'_0$  denote the zero of  $\Delta(t, t_0)$  next greater than  $t_0$  (the "conjugate to  $t_0$ "), the necessary and sufficient condition for the validity of the relation  $\delta^2 I > 0$  is that

$$t_1 < t'_0. \quad (3)$$

That this condition ("Jacobi's condition") is necessary is easily seen.<sup>6</sup> For the integral  $\delta^2 I$  can be thrown, by an integration by parts, into the form

$$\delta^2 I = \int_{t_0}^{t_1} w \Psi(w) dt, \quad (4)$$

so that, on account of (b),

$$\delta^2 I = \int_{t_0}^{t_1} w (\Psi(w) - \mu T) dt, \quad (5)$$

$\mu$  being any constant. And if  $t'_0 \equiv t_1$ , we could make  $\delta^2 I = 0$  by choosing

$$\begin{aligned} \mu &= M(t'_0), \\ w &= M(t'_0) V(t) - N(t'_0) U(t) \quad \text{on } (t_0, t'_0), \\ w &= 0 \quad \text{on } (t'_0, t_1); \end{aligned}$$

this function  $w$  fulfils the conditions (a), (b), (c), and, besides, it satisfies the differential equation

$$\Psi(w) = M(t'_0) T,$$

and therefore it makes  $\delta^2 I = 0$ .<sup>7</sup>

It is less evident that the condition (3) is also sufficient, and to show this is the object of the present note.

<sup>5</sup> It is well known that the general integral of

$$\Psi(w) = \mu T,$$

where  $\mu$  is any constant, can be derived from the general integral of the differential equation of the extremals by differentiation with respect to the constants of integration and the isoperimetric constant  $\lambda$ . See HORMANN, *Dissertation*, Göttingen, 1887, and KNESER, *Mathematische Annalen*, Vol. LV (1901), p. 93.

<sup>6</sup> Compare HORMANN, *loc. cit.*, and KNESER, *loc. cit.*, the proof had already been given, in slightly different form, in Weierstrass's lectures on "The Calculus of Variations," of 1872.

<sup>7</sup>  $\delta^2 I$  can even be made  $< 0$ , as shown by KNESER, *loc. cit.*; compare also BOLZA, "Zur zweiten Variation bei isoperimetrischen Problemen," in one of the forthcoming numbers of the *Mathematische Annalen*.

## §2. PROOF OF THE SUFFICIENCY OF JACOBI'S CONDITION\*

Our proof is based upon an extension of the lemma concerning the differential expression  $\Psi(u)$  by which Jacobi proves the analogous theorem for the unconditioned problem, viz.:

If  $u$  be any integral of the differential equation

$$\Psi(u) = 0,$$

then for every function  $p$  admitting first and second derivatives the relation

$$(pu)\Psi(pu) = H_1(p'u)^2 - \frac{d}{dt}H_1pp'u^2$$

holds, accents denoting differentiation with respect to  $t$ .

To obtain the desired extension of this lemma, let  $u, v$  be integrals of the differential equations

$$\Psi(u) = 0, \quad \Psi(v) = T, \quad (6)$$

respectively, both vanishing in a point  $t = \tau_0$ :

$$u(\tau_0) = 0, \quad v(\tau_0) = 0; \quad (7)$$

let further  $p, q$  be two arbitrary functions of  $t$  admitting first and second derivatives, and denote for shortness

$$\omega = pu + qv.$$

Then

$$\omega\Psi(\omega) = (pu + qv)qT - H_1(pu + qv)(p'u' + q'v') - (pu + qv)\frac{d}{dt}H_1(p'u + q'v),$$

which may easily be written

$$\omega\Psi(\omega) = H_1(p'u + q'v)^2 - H_1(pq' - p'q)(uv' - u'v) + (pu + qv)qT - \frac{d}{dt}H_1(pu + qv)(p'u + q'v).$$

Further, if we introduce

$$m = \int_{\tau_0}^t u T dt, \quad n = \int_{\tau_0}^t v T dt,$$

then

$$(pu + qv)qT = (pm' + qn')q = \frac{d}{dt}(pm + qn)q - (p'm + q'n)q - (pm + qn)q'.$$

But from (6) and (7) follows that<sup>†</sup>

$$H_1(uv' - u'v) = -m;$$

\* Inasmuch as the isoperimetric problem here considered is a special case of the general problem: to minimize an integral of the form

$$I = \int_{t_0}^{t_1} F(t; x_1, x_2, \dots, x_n; x_1', x_2', \dots, x_n') dt,$$

the unknown functions  $x_1, x_2, \dots, x_n$  being connected by a number of relations of the form:

$$\phi_a(t; x_1, x_2, \dots, x_n; x_1', x_2', \dots, x_n') = 0 \quad (a = 1, 2, \dots, m),$$

the proof might be derived by specialization from Mayer's researches on the second variation for the general problem (Crelle's *Journal*, Vol. LXIX (1868), p. 238, and *Mathematische Annalen*, Vol. XIII (1878), p. 53; compare also C. JORDAN, *Cours d'Analyse*, T. III, Nos. 373-94, and VON ESCHERICH, *Sitzungsberichte der Wiener Academie*, Vol. CVII, II, (1896), pp. 1191, 1268, 1384). A proof thus obtained would, however, be much more complicated than the direct proof given in the text.

<sup>†</sup> Compare KNESSLER, *loc. cit.*, equation (22).

## 6 THE SECOND VARIATION IN THE SO-CALLED ISOPERIMETRIC PROBLEMS

hence we obtain

$$(pu + qv)\Psi(pu + qv) = H_1(p'u + q'v)^2 - 2q(p'm + q'n) - \frac{d}{dt}[H_1(pu + qv)(p'u + q'v) - (pm + qn)q], \quad (8)$$

which is the desired extension of Jacobi's lemma.

Since

$$H_1\omega'^2 + H_2\omega^2 = \omega\Psi(\omega) + \frac{d}{dt}(H_1\omega\omega'),$$

we further derive from (8):

$$H_1\omega'^2 + H_2\omega^2 = H_1(p'u + q'v)^2 - 2q(p'm + q'n) + \frac{d}{dt}[H_1(pu + qv)(pu' + qv') + (pm + qn)q]. \quad (9)$$

In the latter formula the second derivatives of  $p$  and  $q$  do not occur; hence it can be inferred that (9) holds even for functions  $p, q$  admitting first but not second derivatives.

We now proceed to prove the sufficiency<sup>10</sup> of Jacobi's condition. We suppose then that

$$t_1 < t'_0,$$

so that

$$\Delta(t, t_0) \neq 0 \quad \text{for every } t, \quad t_0 < t \leq t_1.$$

Now choose  $\tau_0 < t_0$ , but so near to it that  $H_1, H_2, T$  remain regular in the enlarged interval  $(\tau_0, t_1)$  and that

$$\Delta(t, \tau_0) \neq 0, \quad (10)$$

for every  $t, t_0 \leq t \leq t_1$ ; such a choice of  $\tau_0$  is always possible.<sup>11</sup>

Let  $w$  be any function of  $t$  satisfying the conditions (a), (b), (c) for the interval  $(t_0, t_1)$  and  $\equiv 0$  on  $(\tau_0, t_0)$ ; then define the two functions  $p, q$  by the two equations

$$\begin{aligned} pu + qv &= w, \\ p'm + q'n &= 0, \end{aligned} \quad (11)$$

with the initial conditions

$$p(t_0) = 0, \quad q(t_0) = 0. \quad (12)$$

But from (11) it follows that

$$\frac{d}{dt}(pm + qn) = (pu + qv)T = wT,$$

hence, integrating and remembering (12), we obtain

$$pm + qn = \int_{t_0}^t wT dt. \quad (13)$$

<sup>10</sup> It is hardly necessary to say that we are always speaking of sufficiency for a permanent sign of  $\delta^2 I$ , not of sufficiency for a minimum.

<sup>11</sup> For a proof see C. JORDAN, *Cours d'Analyse*, T. III, No. 393.

We thus obtain for the determination of  $p, q$  two linear equations whose determinant is

$$mv - nu = \Delta(t, \tau_0),$$

and therefore  $\neq 0$  on  $(t_0, t_1)$  according to (10); hence  $p$  and  $q$  are continuous on  $(t_0, t_1)$  and with respect to their derivatives of the same character as  $w$ .

For the same reason  $p$  and  $q$  both vanish, not only in  $t_0$ , but also in  $t_1$ , since, according to (a) and (b),  $w$  and

$$\int_{t_0}^t w T dt$$

vanish in  $t_0$ , and in  $t_1$ . Using these functions  $p$  and  $q$  in the transformation (9) and integrating<sup>12</sup> between the limits  $t_0$  and  $t_1$ , we obtain the *final result*:

$$\delta^2 I = \int_{t_0}^{t_1} H_1(p'u + q'v)^2 dt. \quad (14)$$

*This proves that indeed*

$$\delta^2 I > 0$$

for all functions  $w$  satisfying the conditions (a), (b), (c), provided  $t_1 < t'_0$ . For  $\delta^2 I > 0$  unless  $p'u + q'v$  were identically zero; but then it would follow from  $p'm + q'n = 0$  and (10) that  $p'$  and  $q'$  must vanish identically, and therefore also  $p$  and  $q$  themselves, since they vanish in  $t_0$  and are continuous; but this is against the assumption that  $w$  does not vanish identically.

<sup>12</sup> If  $w$ , and accordingly  $p'$  and  $q'$ , have discontinuities of the kind admitted by our assumptions, the integral would have to be broken up into a sum of integrals taken

over the subintervals; but as  $p$  and  $q$  themselves are continuous, formula (9) shows that (14) remains true also in this case.





**TERNARY ORTHOGONAL GROUP IN A GENERAL  
FIELD**



# TERNARY ORTHOGONAL GROUP IN A GENERAL FIELD

LEONARD EUGENE DICKSON

1. FOR two classes of fields (realms of rationality) the theory of linear groups has been investigated at great length: for the continuous real and complex fields by Lie and his followers; for finite fields by Galois, Jordan, and many recent writers.<sup>1</sup> The theory of linear groups in an arbitrary field includes these two important classes as the two extreme cases, between which lies a region the investigation of which has yet found but a beginning.<sup>2</sup> A new problem in this direction is the determination of the structure of the ternary orthogonal group in an arbitrary field, the results being known for the above two classes of fields.<sup>3</sup> The investigation, here first published, was made toward the end of the year 1900.

2. A ternary linear transformation, with coefficients in a field  $F$ ,

$$A: \quad \xi'_j = a_{j1}\xi_1 + a_{j2}\xi_2 + a_{j3}\xi_3 \quad (j=1, 2, 3),$$

is said to be *orthogonal* if it leaves formally and absolutely invariant

$$\xi_1^2 + \xi_2^2 + \xi_3^2.$$

The totality of orthogonal transformations  $A$  of non-vanishing determinant constitutes a group, called the *ternary orthogonal group*. Set  $i = \sqrt{-1}$  and denote by  $F(i)$  the field obtained by adjoining the quantity  $i$  to the field  $F$ . The old and new fields may or may not be identical. Introduce the new variables

$$\eta_1 \equiv -i\xi_1, \quad \eta_2 \equiv \xi_2 - i\xi_3, \quad \eta_3 \equiv \xi_2 + i\xi_3,$$

the case in which  $F$  has a modulus  $p=2$  being here excluded. Then

$$\eta_1\eta_3 - \eta_2^2 \equiv \xi_1^2 + \xi_2^2 + \xi_3^2. \quad (1)$$

Then  $A$  becomes a transformation  $A_1$  on the variables  $\eta_1, \eta_2, \eta_3$ ,

$$A_1: \quad \eta'_j = a'_{j1}\eta_1 + a'_{j2}\eta_2 + a'_{j3}\eta_3 \quad (j=1, 2, 3),$$

in which the  $a'_{jk}$  belong to  $F(i)$  and have values determined later. Also, by (1),

$$\eta'_1\eta'_3 - \eta_1'^2 \equiv \eta_1\eta_3 - \eta_2^2. \quad (2)$$

Removing for a moment the limitation that  $\eta_1, \eta_2, \eta_3$  are independent variables, we set

$$\eta_1 = xy, \quad \eta_2 = x^2, \quad \eta_3 = y^2.$$

<sup>1</sup> JORDAN, *Traité des substitutions*, 1870; DICKSON, *Linear Groups*, 1901.

<sup>2</sup> Compare the author's articles in the *Transactions of the American Mathematical Society*, Vol. II (1901), pp. 363-94; *Proceedings of the London Mathematical Society*, Vol.

XXXIV (1902), pp. 185-205; *Quarterly Journal of Mathematics*, Vol. XXXIII (1901), pp. 145-73.

<sup>3</sup> For the continuous fields see WEBER, *Algebra*, Vol. II, pp. 184-94 (1st ed.); pp. 244-54 (2d ed.). For finite fields see *Linear Groups*, p. 164.

Then  $\eta'_1, \eta'_2, \eta'_3$  become binary quadratic forms in  $x$  and  $y$  with coefficients in  $F(i)$ . In view of (2),  $\eta'_2 \eta'_3 \equiv \eta_1'^2$ . Let  $\eta'_2, \eta'_3$  be decomposed into linear functions of  $x, y$ . If either factor of  $\eta'_2$  occurs in  $\eta'_3$ , then both factors of  $\eta'_2$  occur in  $\eta'_3$ , since the product of the latter is the square of  $\eta_1'$ . This is impossible, since the coefficients of  $\eta'_2$  and  $\eta'_3$  would be proportional, making  $|a'_{jk}| = 0$ . Hence the factors of  $\eta'_2$  are distinct from those of  $\eta'_3$  and therefore the factors of each are equal. Now, the double factor of a quadratic form  $\eta'_j$  may be found by a rational process, so that its coefficients may be assumed to belong to the field  $F(i)$ . We may therefore set<sup>4</sup>

$$\eta'_2 = \kappa(ax + \beta y)^2, \quad \eta'_3 = \kappa(\gamma x + \delta y)^2, \quad \eta_1' = \kappa(ax + \beta y)(\gamma x + \delta y),$$

where  $\kappa, a, \beta, \gamma, \delta$  belong to the field  $F(i)$ . Expanding, we get

$$A_1: \begin{array}{l} \eta_1' = \\ \eta_2' = \\ \eta_3' = \end{array} \begin{array}{ccc} \eta_1 & \eta_2 & \eta_3 \\ \hline \kappa(a\delta + \beta\gamma) & \kappa a\gamma & \kappa\beta\delta \\ 2\kappa a\beta & \kappa a^2 & \kappa\beta^2 \\ 2\kappa\gamma\delta & \kappa\gamma^2 & \kappa\delta^2 \end{array}$$

The coefficient of  $\eta_2 \eta_3$  in  $\eta_2' \eta_3' - \eta_1'^2$  must be unity by (2). Hence

$$\kappa^2(a\delta - \beta\gamma)^2 = 1.$$

The determinant of  $A_1$  is seen to equal

$$\kappa^2(a\delta - \beta\gamma)^2.$$

The determinant of an orthogonal transformation, and hence that of  $A_1$ , is  $\pm 1$ . We limit the discussion to transformations of determinant  $+1$  and denote by  $O(3, F)$  the resulting subgroup of the orthogonal group. Then

$$\kappa(a\delta - \beta\gamma) = 1. \quad (3)$$

Giving to the transformation  $A_1$ , subject to (3), the notation

$$\begin{bmatrix} a & \beta \\ \gamma & \delta \end{bmatrix}_\kappa \equiv \begin{bmatrix} \tau a & \tau \beta \\ \tau \gamma & \tau \delta \end{bmatrix}_{\kappa\tau^{-1}} \quad \kappa(a\delta - \beta\gamma) = 1, \quad (4)$$

we readily verify the formula of composition

$$\begin{bmatrix} a & \beta \\ \gamma & \delta \end{bmatrix}_\kappa \begin{bmatrix} a' & \beta' \\ \gamma' & \delta' \end{bmatrix}_{\kappa'} = \begin{bmatrix} a'a + \beta'\gamma & a'\beta + \beta'\delta \\ \gamma'a + \delta'\gamma & \gamma'\beta + \delta'\delta \end{bmatrix}_{\kappa\kappa'}. \quad (5)$$

Written in terms of the initial variables  $\xi_1, \xi_2, \xi_3$ ,  $A_1$  takes the form

$$A': \begin{array}{l} \xi_1' = \\ \xi_2' = \\ \xi_3' = \end{array} \begin{array}{ccc} \xi_1 & \xi_2 & \xi_3 \\ \hline \kappa(a\delta + \beta\gamma) & \kappa i(a\gamma + \beta\delta) & \kappa(a\gamma - \beta\delta) \\ -\kappa i(a\beta + \gamma\delta) & \frac{1}{2}\kappa(a^2 + \beta^2 + \gamma^2 + \delta^2) & \frac{1}{2}\kappa i(-a^2 + \beta^2 - \gamma^2 + \delta^2) \\ \kappa(a\beta - \gamma\delta) & \frac{1}{2}\kappa i(a^2 + \beta^2 - \gamma^2 - \delta^2) & \frac{1}{2}\kappa(a^2 - \beta^2 - \gamma^2 + \delta^2) \end{array}.$$

In view of (3),  $A'$  has determinant unity and leaves  $\xi_1^2 + \xi_2^2 + \xi_3^2$  invariant.

<sup>4</sup> If initially the constant factor in  $\eta_1$  were  $\kappa_1$ , then  $\kappa\kappa_1$  is a square, say  $\lambda^2$ , in the field. Then  $\kappa_1 = \kappa \left(\frac{\lambda}{\kappa}\right)^2$ , so that  $\frac{\lambda}{\kappa}$  may be absorbed into the constants  $\gamma, \delta$ .

3. Suppose first that  $i$  belongs to  $F$ , so that  $F(i) \equiv F$ , and  $-1$  is a square in the field  $F$ . The group  $O(3, F)$  is then composed of the totality of transformations  $A'$  in which  $\kappa, a, \beta, \gamma, \delta$  take all sets of solutions in  $F$  of equation (3). This group is evidently holodrically isomorphic with the group of transformations  $A_1$ . Now (4) is the identity if, and only if,

$$\beta = \gamma = 0, \quad a = \delta = \pm \kappa^{-\frac{1}{2}}.$$

Also the formula of composition (5) is identical with that for the transformations

$$z' = \frac{az + \beta}{\gamma z + \delta} \quad (a\delta - \beta\gamma \neq 0). \quad (6)$$

Hence  $O(3, F)$  is holodrically isomorphic with the group of all linear fractional transformations (6). The latter has an invariant subgroup formed by the transformations for which  $a\delta - \beta\gamma$  is a square (which may be taken to be unity). But the group of all linear fractional transformations of determinant unity is a simple group  $LF(2, F)$  in every field  $F$  (*Transactions, loc. cit.*, p. 367). Hence the transformations  $A'$  in which  $\kappa$  is a square in  $F$  form a simple group  $O'(3, F)$ , an invariant subgroup of  $O(3, F)$ . These two groups are identical when  $F$  is the field  $C$  of all complex numbers, since every such number is the square of a complex number.

To determine the generators of  $O(3, F)$ , consider its transformations

$$O_{\frac{1}{2}, \frac{1}{2}}^{\lambda, \mu}: \quad \xi'_j = \lambda \xi_j + \mu \xi_k, \quad \xi'_k = -\mu \xi_j + \lambda \xi_k \quad (\lambda^2 + \mu^2 = 1). \quad (7)$$

If  $F$  has a modulus  $p$ , we assume<sup>5</sup> here that  $p > 5$ . Then  $F$  contains elements  $\rho$  for which  $\rho^4 \neq 1$ , so that the elements

$$\tau = \frac{1}{2}(\rho^{-1} + \rho), \quad \beta = \frac{1}{2}(\rho^{-1} - \rho)$$

are different from zero and satisfy  $1 + \beta^2 = \tau^2$ . To the product

$$O_{\frac{1}{2}, \frac{1}{2}}^{\lambda, \mu} \quad O_{\frac{1}{2}, \frac{1}{2}}^{\tau^{-1}, \beta\tau^{-1}} \quad O_{\frac{1}{2}, \frac{1}{2}}^{\beta, -\tau} \quad O_{\frac{1}{2}, \frac{1}{2}}^{\frac{1}{2}(\tau^{-1} + \tau), \frac{1}{2}(\tau^{-1} - \tau)}$$

corresponds the transformation  $\begin{bmatrix} 1 & \beta \\ 0 & 1 \end{bmatrix}_1$ . The latter is transformed into  $\begin{bmatrix} 1 & \beta\sigma^2 \\ 0 & 1 \end{bmatrix}_1$  by  $\begin{bmatrix} \sigma & 0 \\ 0 & \sigma^{-1} \end{bmatrix}_1$ , which corresponds to a certain  $O_{2,3}$ . Also

$$\begin{bmatrix} 1 & \beta\sigma_1^2 \\ 0 & 1 \end{bmatrix}_1 \begin{bmatrix} 1 & \beta\sigma_2^2 \\ 0 & 1 \end{bmatrix}_1 = \begin{bmatrix} 1 & \beta\sigma_1^2 \\ 0 & 1 \end{bmatrix}_1 \begin{bmatrix} 1 & -\beta\sigma_2^2 \\ 0 & 1 \end{bmatrix}_1 = \begin{bmatrix} 1 & \beta(\sigma_1^2 - \sigma_2^2) \\ 0 & 1 \end{bmatrix}_1.$$

For  $\sigma_1 = \frac{1}{2}(\lambda\beta^{-1} + 1)$ ,  $\sigma_2 = \frac{1}{2}(\lambda\beta^{-1} - 1)$ , the last transformation becomes

$$\begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix}_1. \quad (8)$$

Now  $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}_1$ , which corresponds to  $O_{\frac{1}{2}, \frac{1}{2}}^{-1, 0}$ , transforms (8) into

$$\begin{bmatrix} 1 & 0 \\ -\lambda & 1 \end{bmatrix}_1. \quad (9)$$

<sup>5</sup>The case  $p = 5$  is exceptional. Compare *Linear Groups*, p. 166.

But the types (8) and (9) generate every transformation

$$\begin{bmatrix} a & \beta \\ \gamma & \delta \end{bmatrix}. \quad (10)$$

Expressing the transformation  $O_{\frac{1}{2}, \frac{1}{2}}^{a, b}$  in terms of the variables  $\eta_1, \eta_2, \eta_3$ , we get

$$\eta'_1 = \eta_1, \quad \eta'_2 = (a + bi)\eta_2, \quad \eta'_3 = (a - bi)\eta_3 \quad (a^2 + b^2 = 1).$$

We can make  $a + bi = \nu$ , where  $\nu$  is arbitrary in  $F$ , by taking

$$a = \frac{1}{2}(\nu^{-1} + \nu), \quad b = \frac{i}{2}(\nu^{-1} - \nu).$$

Choosing  $\nu = \kappa a^2$ , we reach the transformation

$$\begin{bmatrix} a & 0 \\ 0 & \kappa^{-1} a^{-1} \end{bmatrix}.$$

The latter, together with (10), generate all the transformations (4). We may state the

**Theorem:** *For  $-1$  a square in the field  $F$ , not having a modulus  $\equiv 5$ , the orthogonal group  $O(3, F)$  is generated by the transformations  $O_{\frac{1}{2}, \frac{1}{2}}^{a, \kappa}$ . Any transformation of  $O(3, F)$  may be expressed as a product  $W O_{\frac{1}{2}, \frac{1}{2}}^{a, \kappa}$ , where  $W$  belongs to the group  $O'(3, F)$  of transformations  $A'$  with  $\kappa = 1$ , and  $O_{\frac{1}{2}, \frac{1}{2}}^{a, \kappa}$  is of the form  $A'$  with  $\kappa$  a not-square.*

4. Suppose next that  $-1$  is a not-square in  $F$ . We seek the necessary and sufficient conditions under which the coefficients of  $A'$  shall belong to  $F$ . Since  $\kappa(a\gamma - \beta\delta)$  and  $\kappa i(a\gamma + \beta\delta)$  are to belong to  $F$ ,  $\kappa a\gamma$  must be conjugate with  $-\kappa\beta\delta$  with respect to  $F$ . Since  $\kappa(a\beta - \gamma\delta)$  and  $-\kappa i(a\beta + \gamma\delta)$  are to belong to  $F$ ,  $\kappa a\beta$  must be conjugate with  $-\kappa\gamma\delta$ . Since  $\kappa(a^2 + \delta^2)$  and  $\kappa i(a^2 - \delta^2)$  are to belong to  $F$ ,  $\kappa a^2$  must be conjugate with  $\kappa\delta^2$ . Similarly,  $\kappa\beta^2$  must be conjugate with  $\kappa\gamma^2$ . Since  $\kappa(a\delta + \beta\gamma)$  and  $\kappa(a\delta - \beta\gamma) = 1$  are to belong to  $F$ ,  $\kappa a\delta$  and  $\kappa\beta\gamma$  must belong to  $F$ . If  $a, \beta, \gamma, \delta$  are not zero, we find by division that  $\frac{\gamma}{\beta}$  must be conjugate with  $\frac{\beta}{\gamma}$ ,  $\frac{a}{\delta}$  with  $\frac{\delta}{a}$ ,  $\frac{\gamma}{a}$  with  $\frac{-\beta}{\delta}$ . Hence, if  $\bar{x}$  denotes the conjugate of  $x$ ,

$$\beta\bar{\beta} = \gamma\bar{\gamma}, \quad a\bar{a} = \delta\bar{\delta}, \quad \bar{a}\beta = -\bar{\gamma}\delta, \quad a\bar{\beta} = -\gamma\bar{\delta}. \quad (11)$$

These equations (11) hold if any of the elements vanish. Indeed,  $\beta$  and  $\gamma$  both vanish if either vanishes; likewise for  $a$  and  $\delta$ . If, for example,  $\beta = \gamma = 0$ , then

$$\kappa a\delta = 1, \quad \kappa a^2 = \bar{\kappa}\bar{\delta}^2.$$

Hence

$$\frac{a}{\delta} = \frac{\kappa a^2}{\kappa a\delta} = \kappa a^2 = \bar{\kappa}\bar{\delta}^2 = \frac{\bar{\kappa}\bar{\delta}^2}{\bar{\kappa}\bar{a}\bar{\delta}} = \frac{\bar{\delta}}{\bar{a}},$$

so that equations (11) follow. Then

$$\begin{aligned} 1 &= \kappa(a\delta - \beta\gamma)\bar{\kappa}(\bar{a}\bar{\delta} - \bar{\beta}\bar{\gamma}) = \kappa\bar{\kappa}(a\bar{a} + \beta\bar{\beta}), \\ 1 &= \kappa\bar{\kappa}\tau^2, \quad \tau \equiv a\bar{a} + \beta\bar{\beta}. \end{aligned} \quad (12)$$

Then  $\kappa\delta^2 = \bar{\kappa}\bar{a}^2$  and  $\kappa a\gamma = -\bar{\kappa}\bar{\beta}\bar{\delta}$  give respectively

$$\delta = \pm \bar{\kappa}\tau\bar{a}, \quad \gamma = \mp \bar{\kappa}\tau\bar{\beta}.$$

Hence (3) is satisfied if, and only if, the upper signs hold.

Inversely, if  $\delta, \gamma$  have these values, the conditions that the coefficients of  $A'$  shall all belong to  $F$  are seen to be satisfied.

It is convenient to introduce in place of  $\kappa$  a new quantity  $\sigma = \bar{\kappa}\tau$ . Then

$$\delta = \sigma\bar{a}, \quad \gamma = -\sigma\bar{\beta}, \quad \sigma\bar{\sigma} = 1, \quad \tau \equiv a\bar{a} + \beta\bar{\beta}. \quad (13)$$

If we replace  $a$  by  $\mu a$ ,  $\beta$  by  $\mu\beta$ ,  $\kappa$  by  $\kappa\mu^{-2}$ , where  $\mu$  is any element  $\neq 0$  of the field  $F(i)$ , then  $\tau$  is replaced by  $\mu\bar{\mu}\tau$ ,  $\sigma$  by  $\frac{\sigma\mu}{\mu}$ ,  $\gamma$  by  $\mu\gamma$ ,  $\delta$  by  $\mu\delta$ . Hence, in view of (4), we may give to  $A'$  the notation

$$\begin{pmatrix} a & \beta \\ -\sigma\bar{\beta} & \sigma\bar{a} \end{pmatrix} \equiv \begin{pmatrix} \mu a & \mu\beta \\ -\mu\sigma\bar{\beta} & \mu\sigma\bar{a} \end{pmatrix} \quad (a\bar{a} + \beta\bar{\beta} \neq 0, \sigma\bar{\sigma} = 1). \quad (14)$$

It follows that the group  $O(3, F)$  is holodrically isomorphic with the group of all linear fractional transformations of determinant  $\sigma(a\bar{a} + \beta\bar{\beta}) \neq 0$ ,

$$z' = \frac{az + \beta}{-\sigma\bar{\beta}z + \sigma\bar{a}}. \quad (15)$$

The latter group has as a subgroup the group  $H$  of transformations (15) with  $a\bar{a} + \beta\bar{\beta} = 1$ , whose determinant  $\sigma$  is therefore a root of  $\sigma\bar{\sigma} = 1$ . Now,  $H$  is the group of all fractional (binary) hyperorthogonal\* transformations (*Linear Groups*, p. 132). It evidently has an invariant subgroup composed of all fractional hyperorthogonal transformations of determinant unity:

$$z' = \frac{az + \beta}{-\bar{\beta}z + \bar{a}} \quad (a\bar{a} + \beta\bar{\beta} = 1). \quad (16)$$

If  $F$  be the field  $R$  of all real numbers,  $F(i)$  is the field  $C$  of all complex numbers. Then  $\kappa$  is a square in  $C$  and may be absorbed into the  $a, \beta, \gamma, \delta$  of  $A_1$ , so that we may take  $\kappa = 1$ . Then (13) follows, with  $\sigma = \tau$ . Hence  $\sigma^2 = 1$ . Since  $\sigma = a\bar{a} + \beta\bar{\beta}$  belongs to  $R$  and is positive, we have  $\sigma = 1$ . Hence  $O(3, R)$  is holodrically isomorphic with the group  $H_1$  of all fractional hyperorthogonal transformations (16) with coefficients in  $C$  of determinant unity. This result is in accord with a theorem due to Cayley.<sup>7</sup> For the case of the continuous field  $C$ , the structure of  $H_1$  and of the analogous group on an arbitrary number of variables has been investigated by the writer.<sup>8</sup>

5. Consider the group of hyperorthogonal transformations

$$\begin{pmatrix} a & \beta \\ -\bar{\beta} & \bar{a} \end{pmatrix}, \quad (a\bar{a} + \beta\bar{\beta} = 1). \quad (17)$$

\*The corresponding binary homogeneous transformations

$$z_1' = az_1 + \beta z_2, \quad z_2' = -\sigma\bar{\beta}z_1 + \sigma\bar{a}z_2$$

leave formally invariant the function  $z_1\bar{z}_1 + z_2\bar{z}_2$ .

<sup>7</sup>*Mathematische Annalen*, Vol. XV (1879); cf. KLEIN, *Ikoseder*, p. 34.

<sup>8</sup>*Bulletin of the American Mathematical Society*, Vol. VII (1901) pp. 340-50.



The formula of composition for transformations (17) is

$$\begin{pmatrix} a & \beta \\ -\bar{\beta} & \bar{a} \end{pmatrix} \begin{pmatrix} a' & \beta' \\ -\bar{\beta}' & \bar{a}' \end{pmatrix} = \begin{pmatrix} a'' & \beta'' \\ -\bar{\beta}'' & \bar{a}'' \end{pmatrix}, \quad \begin{cases} a'' = a'a - \beta'\bar{\beta} \\ \beta'' = a'\beta + \beta'\bar{a} \end{cases}.$$

If  $\bar{a} = a$ ,  $\bar{\beta} = \beta$ ,  $\bar{a}' = a'$ ,  $\bar{\beta}' = \beta'$ , then  $\bar{a}'' = a''$ ,  $\bar{\beta}'' = \beta''$ . Hence the group of the transformations (17) contains as a subgroup the binary *orthogonal* group of the transformations  $\begin{pmatrix} a & \beta \\ -\beta & a \end{pmatrix}$ , in which  $a$  and  $\beta$  are any solutions in the field  $F$  of the equation  $a^2 + \beta^2 = 1$ . This is a *commutative* group, since

$$\begin{pmatrix} a & \beta \\ -\beta & a \end{pmatrix} \begin{pmatrix} a' & \beta' \\ -\beta' & a' \end{pmatrix} = \begin{pmatrix} a'' & \beta'' \\ -\beta'' & a'' \end{pmatrix}, \quad \begin{cases} a'' = a'a - \beta'\beta \\ \beta'' = a'\beta + \beta'a \end{cases},$$

the values of  $a''$  and  $\beta''$  being unaltered by the interchange of  $a$  with  $a'$ ,  $\beta$  with  $\beta'$ .

The binary *hyperorthogonal* group in the field  $F(i)$  has as a subgroup the binary *orthogonal commutative* group in the field  $F$ .

6. Consider the transformations (17) when the field  $F$  is the field of all *rational* numbers. Set

$$a = \frac{a}{e} + i\frac{b}{e}, \quad \beta = \frac{c}{e} + i\frac{d}{e} \quad (a, b, c, d, e \text{ all integers}).$$

The condition  $a\bar{a} + \beta\bar{\beta} = 1$  becomes

$$a^2 + b^2 + c^2 + d^2 = e^2. \quad (18)$$

Set  $e = a + n$ . In view of (18),

$$e = \frac{(b^2 + c^2 + d^2 + n^2)}{2n}, \quad a = \frac{(b^2 + c^2 + d^2 - n^2)}{2n}. \quad (19)$$

Inversely, if  $b, c, d, n$  are arbitrary integers not all zero, and if  $e$  and  $n$  are defined by (19), then (18) is satisfied and

$$a = \frac{a + ib}{e} = \frac{b^2 + c^2 + d^2 - n^2 + 2ibn}{b^2 + c^2 + d^2 + n^2} = \frac{c^2 + d^2 + (b + in)^2}{b^2 + c^2 + d^2 + n^2},$$

$$\beta = \frac{c + id}{e} = \frac{2n(c + id)}{b^2 + c^2 + d^2 + n^2}.$$

Set  $\kappa = c + id$ ,  $\lambda = n - ib$ , so that  $\kappa$  and  $\lambda$  are arbitrary elements of  $F(i)$ , not both zero. Then

$$a = \frac{\kappa\bar{\kappa} - \lambda^2}{\kappa\bar{\kappa} + \lambda\bar{\lambda}}, \quad \beta = \frac{\kappa(\lambda + \bar{\lambda})}{\kappa\bar{\kappa} + \lambda\bar{\lambda}}, \quad a\bar{a} + \beta\bar{\beta} = 1.$$

In particular, the transformation (17) taken fractionally becomes

$$z' = \frac{(\lambda^2 - \kappa\bar{\kappa})z - \kappa(\lambda + \bar{\lambda})}{\bar{\kappa}(\lambda + \bar{\lambda})z + \bar{\lambda}^2 - \kappa\bar{\kappa}},$$

and is the *square* of the transformation

$$z' = \frac{\lambda z - \kappa}{\bar{\kappa}z + \bar{\lambda}}.$$

**GROUPS DEFINED FOR A GENERAL FIELD BY  
THE ROTATION GROUPS**



# THE GROUPS DEFINED FOR A GENERAL FIELD BY THE ROTATION GROUPS

LEONARD EUGENE DICKSON

1. THE theory and application of group-determinants and group-characters have been developed at length by Frobenius.<sup>1</sup> The theory was then developed from a very different standpoint by Burnside,<sup>2</sup> the methods of investigation and the results belonging to the theory of continuous groups. More recently the theory has been investigated by the present writer<sup>3</sup> from a more general point of view, a generalization of the theory from a continuous field to an arbitrary field being thus obtained. A further contribution to the subject is made in this paper in the way of an application of the generalized theory to the important rotation groups (the cyclic, dihedron, tetrahedral, octahedral, and icosahedral groups). The methods used are quite elementary and are practically independent of the earlier papers.

2. The cyclic group  $g_k$  of order  $k$  is composed of the operators

$$a_0 = \text{identity}, \quad a_i = a^i \quad (i = 1, 2, \dots, k-1).$$

We form the multiplication table for  $g_k$  with the operators  $a_0, a_1, \dots, a_{k-1}$  in the first row and their inverses in the same order as the left-hand multipliers. The body of the table gives the matrix  $A$ :

$$A \equiv \begin{pmatrix} a_0 & a_1 & a_2 & \dots & a_{k-2} & a_{k-1} \\ a_{k-1} & a_0 & a_1 & \dots & a_{k-3} & a_{k-2} \\ a_{k-2} & a_{k-1} & a_0 & \dots & a_{k-4} & a_{k-3} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_2 & a_3 & a_4 & \dots & a_0 & a_1 \\ a_1 & a_2 & a_3 & \dots & a_{k-1} & a_0 \end{pmatrix}, \quad R \equiv \begin{pmatrix} \rho_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \rho_1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \rho_2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \rho_{k-2} & 0 \\ 0 & 0 & 0 & \dots & 0 & \rho_{k-1} \end{pmatrix}.$$

When  $a_0, a_1, \dots, a_{k-1}$  are given arbitrary values in a given field  $F$ ,  $A$  is called the *group-matrix* of  $g_k$ . Let  $\omega$  be a primitive  $k$ -th root of unity and set

$$\begin{aligned} \rho_i &\equiv a_0 + \omega^i a_1 + \omega^{2i} a_2 + \dots + \omega^{(k-1)i} a_{k-1} \\ \eta_i &\equiv \xi_0 + \omega^{-i} \xi_1 + \omega^{-2i} \xi_2 + \dots + \omega^{-(k-1)i} \xi_{k-1} \end{aligned} \quad (i = 0, 1, \dots, k-1).$$

The determinant of the coefficients of  $a_0, \dots, a_{k-1}$  in  $\rho_0, \dots, \rho_{k-1}$  equals the product of the differences  $\omega^r - \omega^s$  and is not zero unless  $x^k = 1$  has a double root in the field  $F$ . The latter happens<sup>4</sup> only for a field having a modulus which divides  $k$ . Exclud-

<sup>1</sup> *Berliner Sitzungsberichte*, 1896, pp. 985-1021, 1343-82; 1897, pp. 994-1015; 1898, pp. 501-16; 1899, pp. 330-40, 482-500; 1900, pp. 516-34; 1901, pp. 303-15.

<sup>2</sup> "On the Continuous Group That is Defined by Any Given Group of Finite Order," *Proceedings of the London Mathematical Society*, Vol. XXIX (1898), pp. 207-24, 546-65.

<sup>3</sup> "On the Group Defined for Any Given Field by the Multiplication Table of Any Given Finite Group," *Transactions of the American Mathematical Society*, Vol. III (1902), pp. 285-301.

<sup>4</sup> DICKSON, *Linear Groups*, p. 54, Corollary of § 74, the discussion holding for an arbitrary field.

ing this case, we conclude that  $\rho_0, \dots, \rho_{k-1}$  are linearly independent functions of  $a_0, \dots, a_{k-1}$ , and likewise that the  $\eta_i$  are linearly independent functions of the  $\xi_i$ . A transformation on  $\xi_0, \xi_1, \dots, \xi_{k-1}$ , whose coefficients are given by matrix  $A$ , gives rise to the following transformation on  $\eta_0, \eta_1, \dots, \eta_{k-1}$ :

$$\eta'_i = \rho_i \eta_i \quad (i = 0, 1, \dots, k-1). \quad (1)$$

If  $F$  does not have a modulus which divides  $k$ , matrix  $A$  can be transformed into matrix  $R$ . The determinant of  $A$  equals  $\rho_0 \rho_1 \dots \rho_{k-1}$ .

If  $\omega$  belongs to  $F$ , matrix  $R$  belongs to  $F$ . Inversely, if arbitrary values  $\neq 0$  in  $F$  be assigned to  $\rho_0, \dots, \rho_{k-1}$ , then  $a_0, \dots, a_{k-1}$  are uniquely determined in  $F$ . Since the totality of the resulting transformations (1) forms a group, the totality of transformations defined by matrix  $A$  by giving  $a_0, \dots, a_{k-1}$  arbitrary values in  $F$  such that  $|A| \neq 0$  forms a group. The latter is therefore simply isomorphic with the direct product of  $k$  unary linear groups in  $F$ .

If  $\omega$  does not belong to  $F$ , we decompose  $x^k - 1$  into factors belonging to and irreducible in  $F$ . For example, if  $k = 6$ ,

$$x^6 - 1 = (x-1)(x+1)(x^2 - x + 1)(x^2 + x + 1).$$

Assuming that  $F$  does not have modulus 2 or 3, a root of  $x^2 - x + 1 = 0$  may be taken as  $\omega$ , a primitive sixth root of unity. Then  $\bar{\omega} = -\omega^2 = \omega^5$ . Hence  $\rho_1$  and  $\rho_5$ ,  $\rho_2$  and  $\rho_4$ , are conjugate with respect to  $F$ ; likewise the functions  $\eta_1$  and  $\eta_5$ ,  $\eta_2$  and  $\eta_4$ . Let  $\omega$  extend  $F$  to the larger field  $F_2$ . If we assign to  $\rho_0$  and  $\rho_3$  arbitrary values  $\neq 0$  in  $F$ , to  $\rho_1$  and  $\rho_5$  arbitrary conjugate values  $\neq 0$  in  $F_2$ , and to  $\rho_2$  and  $\rho_4$  arbitrary conjugate values  $\neq 0$  in  $F_2$ , it is readily verified that  $a_0, \dots, a_5$  are uniquely determined in  $F$ . The resulting transformations (1) form a group which is simply isomorphic with the direct product of two unary groups in  $F$  and two unary groups in  $F_2$ :

$$(\eta'_0 = \rho_0 \eta_0), \quad (\eta'_3 = \rho_3 \eta_3), \quad (\eta'_1 = \rho_1 \eta_1), \quad (\eta'_5 = \rho_5 \eta_5).$$

Hence the totality of transformations defined by matrix  $A$  by giving  $a_0, \dots, a_{k-1}$  arbitrary values in  $F$  such that  $|A| \neq 0$  forms an isomorphic group.

3. The symmetric group on three letters is formed by the substitutions

$$x = \text{identity}, \quad y = (123), \quad z = (132), \quad u = (12), \quad v = (23), \quad w = (13).$$

Denote by  $A, B, C$  the respective matrices:

$$\begin{pmatrix} x & y & z & u & v & w \\ z & x & y & w & v & u \\ y & z & x & v & u & w \\ u & w & v & x & y & z \\ w & v & u & z & x & y \\ v & u & w & y & z & x \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & -1 & -1 \\ -1 & 1 & 0 & -1 & 1 & 0 \\ -1 & 0 & 1 & -1 & 0 & 1 \\ 1 & 1 & -2 & -1 & -1 & 2 \\ -1 & 2 & -1 & 1 & -2 & 1 \end{pmatrix}, \quad \begin{pmatrix} \lambda & 0 & 0 & 0 & 0 & 0 \\ 0 & \mu & 0 & 0 & 0 & 0 \\ 0 & 0 & \nu & \rho & 0 & 0 \\ 0 & 0 & \sigma & \tau & 0 & 0 \\ 0 & 0 & 0 & 0 & \nu & \rho \\ 0 & 0 & 0 & 0 & \sigma & \tau \end{pmatrix}.$$

The first matrix  $A$  is the body of the left-hand multiplication table of the given group. It will be shown that  $BAB^{-1} = C$ , if  $\lambda, \mu, \dots, \tau$  be defined as follows:

$$\begin{aligned} \lambda &= x + y + z + u + v + w, & \nu &= x - y + v - w, & \rho &= y - z + u - v, \\ \mu &= x + y + z - u - v - w, & \sigma &= z - y + u - w, & \tau &= x - z + w - v. \end{aligned} \quad (2)$$

Let  $x, \dots, v$  be arbitrary elements of a field  $F$  such that  $|A| \neq 0$ , and consider the transformation on  $\xi_1, \dots, \xi_6$  whose coefficients are given by matrix  $A$ . Set

$$\eta_1 = \xi_1 + \xi_4, \quad \eta_2 = \xi_1 - \xi_4, \quad \eta_3 = \xi_2 + \xi_5, \quad \eta_4 = \xi_2 - \xi_5, \quad \eta_5 = \xi_3 + \xi_6, \quad \eta_6 = \xi_3 - \xi_6,$$

of determinant  $-8$ , thereby excluding the case of modulus 2. The transformation on  $\eta_1, \dots, \eta_6$  has the matrix  $\begin{pmatrix} D & O \\ O & E \end{pmatrix}$ , where  $O$  is a matrix of zero elements and

$$D \equiv \begin{pmatrix} x+u & y+w & z+v \\ z+w & x+v & y+u \\ y+v & z+u & x+w \end{pmatrix}, \quad E \equiv \begin{pmatrix} x-u & y-w & z-v \\ z-w & x-v & y-u \\ y-v & z-u & x-w \end{pmatrix}.$$

Excluding the case of modulus 3, we may set

$$\begin{aligned} \zeta_1 &= \eta_1 + \eta_2 + \eta_3, & \zeta_2 &= \eta_2 - \eta_1, & \zeta_3 &= \eta_3 - \eta_1, & \zeta_4 &= \eta_4 + \eta_5 + \eta_6, \\ \zeta_5 &= \eta_4 + \eta_6 - 2\eta_5, & \zeta_6 &= 2\eta_5 - \eta_4 - \eta_6, \end{aligned}$$

the determinant being 27. The transformation on  $\zeta_1, \dots, \zeta_6$  has the matrix  $C$  above. The expressions for  $\zeta_1, \dots, \zeta_6$  in terms of  $\xi_1, \dots, \xi_6$  have the matrix of coefficients  $B$  of determinant  $-8 \cdot 27 \equiv -6^3$ . Now, the six functions (2) are linearly independent. Hence the totality of matrices  $A$  may be derived by transformation from the totality of matrices  $C$  where  $\lambda, \mu, \dots, \tau$  take arbitrary values in  $F$  such that  $\lambda\mu(\nu\tau - \rho\sigma)^2 \neq 0$ . Since the latter totality defines a group, the former defines a simply isomorphic group. We may therefore state the theorem:

*The totality of transformations defined by matrix  $A$  by giving  $x, \dots, v$  arbitrary values in  $F$  such that  $|A| \neq 0$  forms a group simply isomorphic with the product of two general unary groups in  $F$  and a general binary group in  $F$ , the cases of modulus 2 and 3 being excluded. The determinant of  $A$  has two distinct linear factors and two equal irreducible quadratic factors.*

#### THE GENERAL DIHEDRON GROUP

4. The dihedron group  $g_{2k}$  is generated by operators  $\alpha, \beta$  for which

$$\alpha^k = \alpha_0 = \text{identity}, \quad \beta^2 = \alpha_0, \quad \alpha\beta = \beta\alpha^{-1}.$$

Set  $\beta_0 = \beta$ ,  $\alpha_i = \alpha^i$ ,  $\beta_i = \beta\alpha^i$  ( $i=1, \dots, k-1$ ). The body of the left-hand multiplication table for  $g_{2k}$  gives the matrix

$$\begin{pmatrix} A & B \\ B & A \end{pmatrix}, \quad (3)$$

where  $A$  is defined in § 2 and  $B$  is defined as follows:

$$B \equiv \begin{pmatrix} \beta_0 & \beta_1 & \beta_2 & \cdots & \beta_{k-2} & \beta_{k-1} \\ \beta_1 & \beta_2 & \beta_3 & \cdots & \beta_{k-1} & \beta_0 \\ \beta_2 & \beta_3 & \beta_4 & \cdots & \beta_0 & \beta_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \beta_{k-1} & \beta_0 & \beta_1 & \cdots & \beta_{k-2} & \beta_{k-1} \end{pmatrix}.$$

Besides the functions  $\rho_i, \eta_i$  of § 2, we consider the following:

$$\begin{aligned} \sigma_i &\equiv \beta_0 + \omega^{-i} \beta_1 + \omega^{-2i} \beta_2 + \cdots + \omega^{-(k-2)i} \beta_{k-2} + \omega^{-(k-1)i} \beta_{k-1}, \\ \zeta_i &\equiv \xi_k + \omega^i \xi_{k+1} + \omega^{2i} \xi_{k+2} + \cdots + \omega^{(k-2)i} \xi_{2k-2} + \omega^{(k-1)i} \xi_{2k-1} \end{aligned} \quad (i=0, 1, \dots, k-1).$$

A transformation on  $\xi_0, \xi_1, \dots, \xi_{2k-1}$ , whose coefficients are given by the matrix (3), gives rise to the following transformation on the independent functions  $\eta_i, \zeta_i$ :

$$T: \quad \eta'_i = \rho_i \eta_i + \sigma_i \zeta_i, \quad \zeta'_i = \sigma_{k-i} \eta_i + \rho_{k-i} \zeta_i \quad (i=0, 1, \dots, k-1).$$

This transformation may be expressed as a product  $T_0 T_1 \cdots T_{k-1}$ , where  $T_i$  alters only  $\eta_j$  and  $\zeta_j$ , replacing them by the same functions that  $T$  does. Now

$$(\eta_j \zeta_{k-j}) (\zeta_j \eta_{k-j})$$

transforms  $T_j$  into  $T_{k-j}$ . The latter are identical for a value of  $j < k$  if, and only if,  $j=0$  or else  $j=\frac{k}{2}$  with  $k$  even. Indeed,  $\omega^j = \omega^{k-j}$  requires that  $\omega^{2j} = 1$ , whence  $j=0$  if  $k$  is odd,  $j=0$  or  $\frac{k}{2}$  if  $k$  is even. But

$$T_0: \quad \eta'_0 = \rho_0 \eta_0 + \sigma_0 \zeta_0, \quad \zeta'_0 = \sigma_0 \eta_0 + \rho_0 \zeta_0$$

gives rise to the transformation

$$x'_0 = (\rho_0 + \sigma_0) x_0, \quad y'_0 = (\rho_0 - \sigma_0) y_0$$

upon the variables  $x_0 \equiv \eta_0 + \zeta_0, y_0 \equiv \eta_0 - \zeta_0$ . Likewise, if  $k=2\kappa$ ,  $T_\kappa$  gives rise to the transformation

$$x'_\kappa = (\rho_\kappa + \sigma_\kappa) x_\kappa, \quad y'_\kappa = (\rho_\kappa - \sigma_\kappa) y_\kappa$$

upon the variables  $x_\kappa \equiv \eta_\kappa + \zeta_\kappa, y_\kappa \equiv \eta_\kappa - \zeta_\kappa$ . Noting that  $\omega^\kappa = -1$ , we have

$$\rho_\kappa = \alpha_0 - \alpha_1 + \alpha_2 - \alpha_3 + \cdots - \alpha_{k-1}, \quad \sigma_\kappa = \beta_0 - \beta_1 + \beta_2 - \beta_3 + \cdots - \beta_{k-1}.$$

According as  $k$  is odd or even ( $k=2\kappa$ ), matrix (3) can be transformed into the first or the second of the following matrices:

$$\begin{pmatrix} \rho_0 + \sigma_0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \rho_0 - \sigma_0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & M_1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & M_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & M_{\frac{(k-1)}{2}} \end{pmatrix}, \quad \begin{pmatrix} \rho_0 + \sigma_0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \rho_0 - \sigma_0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & M_1 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & M_{\kappa-1} & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & \rho_\kappa + \sigma_\kappa & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \rho_\kappa - \sigma_\kappa \end{pmatrix}.$$

where  $M_i \equiv \begin{pmatrix} \rho_i & \sigma_i \\ \sigma_{k-i} & \rho_{k-i} \end{pmatrix}$ . The two or four linear functions  $\rho_0 \pm \sigma_0$ ,  $\rho_\pi \pm \sigma_\pi$  of  $a_0, \dots, a_{k-1}$ ,  $\beta_0, \dots, \beta_{k-1}$  have integral coefficients. The determinants of the matrices  $M_i$  are distinct irreducible quadratic functions of the  $a$ 's and  $\beta$ 's.

These results are in accord with the general theory. For  $k$  odd, the  $2k$  operators fall into the  $\frac{1}{2}(k+3)$  distinct sets of conjugates

$$\{a_0\}, \{a_1, a_{k-1}\}, \{a_2, a_{k-2}\}, \dots, \{a_{\frac{k-1}{2}}, a_{\frac{k+1}{2}}\}, \{\beta_0, \beta_1, \dots, \beta_{k-1}\},$$

and the commutator-group is the cyclic subgroup of order  $k$ . For  $k$  even ( $k=2\kappa$ ), the  $2k$  operators fall into the  $\kappa+3$  distinct sets of conjugates

$$\{a_0\}, \{a_\kappa\}, \{a_1, a_{k-1}\}, \dots, \{a_{\kappa-1}, a_{\kappa+1}\}, \{\beta_0, \beta_2, \beta_4, \dots, \beta_{k-2}\}, \{\beta_1, \beta_3, \beta_5, \dots, \beta_{k-1}\},$$

and the commutator-group is the cyclic group of order  $\kappa$  generated by  $a^2$ .

#### THE ALTERNATING AND SYMMETRIC GROUPS ON FOUR LETTERS

5. These groups  $g_{12}$  and  $g_{24}$  have as an invariant subgroup the four-group formed of the substitutions

$$x_1 = \text{identity}, \quad x_2 = (12)(34), \quad x_3 = (13)(24), \quad x_4 = (14)(23).$$

The eight remaining substitutions of  $g_{12}$  are

$$\begin{aligned} y_1 &= (123), & y_2 &= y_1 x_2 = (243), & y_3 &= y_1 x_3 = (142), & y_4 &= y_1 x_4 = (134), \\ z_1 &= y_1^2 = (132), & z_2 &= y_2^2 = (234), & z_3 &= y_3^2 = (124), & z_4 &= y_4^2 = (143). \end{aligned}$$

We form the multiplication table of  $g_{12}$  with  $x_1, x_2, x_3, x_4, y_1, y_2, y_3, y_4, z_1, z_2, z_3, z_4$  in the first row and their inverses in the same order as the left-hand multipliers. The body of the table gives the group-matrix

$$\begin{pmatrix} (x) & (y) & (z) \\ (z) & (x) & (y) \\ (y) & (z) & (x) \end{pmatrix}, \quad (4)$$

where

$$(x) \equiv \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2 & x_1 & x_4 & x_3 \\ x_3 & x_4 & x_1 & x_2 \\ x_4 & x_3 & x_2 & x_1 \end{pmatrix}, \quad (y) \equiv \begin{pmatrix} y_1 & y_2 & y_3 & y_4 \\ y_4 & y_3 & y_2 & y_1 \\ y_3 & y_1 & y_4 & y_2 \\ y_2 & y_4 & y_1 & y_3 \end{pmatrix}, \quad (z) \equiv \begin{pmatrix} z_1 & z_4 & z_2 & z_3 \\ z_2 & z_3 & z_1 & z_4 \\ z_3 & z_2 & z_4 & z_1 \\ z_4 & z_1 & z_3 & z_2 \end{pmatrix}. \quad (5)$$

Consider the transformation  $T$  on the variables  $\xi_1, \dots, \xi_{12}$  whose matrix of coefficients is (4). For a field  $F$  not having modulus 2 we may introduce as new variables:

$$\begin{aligned} \eta_1 &= \xi_1 + \xi_2 - \xi_3 - \xi_4, & \eta_2 &= \xi_5 - \xi_6 - \xi_7 + \xi_8, & \eta_3 &= \xi_9 - \xi_{10} + \xi_{11} - \xi_{12}, \\ \eta_4 &= \xi_1 - \xi_2 - \xi_3 + \xi_4, & \eta_6 &= \xi_5 - \xi_6 + \xi_7 - \xi_8, & \eta_4 &= \xi_9 + \xi_{10} - \xi_{11} - \xi_{12}, \\ \eta_5 &= \xi_1 - \xi_2 + \xi_3 - \xi_4, & \eta_7 &= \xi_5 + \xi_6 - \xi_7 - \xi_8, & \eta_6 &= \xi_9 - \xi_{10} - \xi_{11} + \xi_{12}, \\ \eta_{10} &= \xi_1 + \xi_2 + \xi_3 + \xi_4, & \eta_{11} &= \xi_5 + \xi_6 + \xi_7 + \xi_8, & \eta_{12} &= \xi_9 + \xi_{10} + \xi_{11} + \xi_{12}. \end{aligned}$$



The matrix of the transformation on  $\eta_1, \dots, \eta_{12}$  is  $M$ , where

$$M \equiv \begin{pmatrix} A & O & O & O \\ O & A & O & O \\ O & O & A & O \\ O & O & O & B \end{pmatrix}, \quad A \equiv \begin{pmatrix} X_1 & Y_1 & Z_1 \\ Z_1 & X_1 & Y_1 \\ Y_1 & Z_1 & X_1 \end{pmatrix}, \quad B \equiv \begin{pmatrix} X_4 & Y_4 & Z_4 \\ Z_4 & X_4 & Y_4 \\ Y_4 & Z_4 & X_4 \end{pmatrix},$$

where  $O$  is a square matrix of order three all of whose elements are zero, and where the following abbreviations are employed:

$$\begin{aligned} X_1 &= x_1 + x_2 - x_3 - x_4, & Y_1 &= y_1 - y_2 - y_3 + y_4, & Z_1 &= z_1 + z_2 - z_3 - z_4, \\ X_2 &= x_1 - x_2 - x_3 + x_4, & Y_2 &= y_1 - y_2 + y_3 - y_4, & Z_2 &= z_1 - z_2 - z_3 + z_4, \\ X_3 &= x_1 - x_2 + x_3 - x_4, & Y_3 &= y_1 + y_2 - y_3 - y_4, & Z_3 &= z_1 - z_2 + z_3 - z_4, \\ X_4 &= x_1 + x_2 + x_3 + x_4, & Y_4 &= y_1 + y_2 + y_3 + y_4, & Z_4 &= z_1 + z_2 + z_3 + z_4. \end{aligned}$$

These 12 functions are evidently independent in  $F$ . Hence the determinant  $|A|$  is an irreducible cubic form in  $F$ . But  $|B| = XYZ$ , where

$$X = X_4 + Y_4 + Z_4, \quad Y = X_4 + \rho Y_4 + \rho^2 Z_4, \quad Z = X_4 + \rho^2 Y_4 + \rho Z_4,$$

$\rho$  being an imaginary cube root of unity. Now the functions

$$\zeta_1 = \eta_{10} + \eta_{11} + \eta_{12}, \quad \zeta_2 = \eta_{10} + \rho^2 \eta_{11} + \rho \eta_{12}, \quad \zeta_3 = \eta_{10} + \rho \eta_{11} + \rho^2 \eta_{12}$$

are multiplied by  $X, Y, Z$ , respectively, on applying the transformation  $T$ . Hence<sup>\*</sup> we obtain a canonical matrix of the form  $M$  with  $B$  replaced by

$$B' \equiv \begin{pmatrix} X & 0 & 0 \\ 0 & Y & 0 \\ 0 & 0 & Z \end{pmatrix}.$$

If  $\rho$  does not belong to the field  $F$ , we apply the transformation

$$\sigma_1 = \eta_{10} + \eta_{11} + \eta_{12}, \quad \sigma_2 = -\eta_{10} + \eta_{11}, \quad \sigma_3 = -\eta_{10} + \eta_{12},$$

of determinant 3. Then  $T$  gives rise to a transformation on  $\sigma_1, \sigma_2, \sigma_3$  with matrix

$$B' \equiv \begin{pmatrix} X_4 + Y_4 + Z_4 & 0 & 0 \\ 0 & X_4 - Y_4 & Y_4 - Z_4 \\ 0 & Z_4 - Y_4 & X_4 - Z_4 \end{pmatrix}.$$

Then matrix  $M$  with  $B'$  in place of  $B$  is the simplest canonical form arising by a transformation with coefficients in  $F$ .

6. The substitutions of  $g_{24}$  not belonging to  $g_{12}$  are

$$\begin{aligned} u_1 &= x_1 u_1 = (12), & u_2 &= x_2 u_1 = (34), & u_3 &= x_3 u_1 = (1324), & u_4 &= x_4 u_1 = (1423), \\ v_1 &= y_1 u_1 = (23), & v_2 &= y_2 u_1 = (1243), & v_3 &= y_3 u_1 = (14), & v_4 &= y_4 u_1 = (1342), \\ w_1 &= z_1 u_1 = (13), & w_2 &= z_2 u_1 = (1234), & w_3 &= z_3 u_1 = (24), & w_4 &= z_4 u_1 = (1432). \end{aligned}$$

<sup>\*</sup> As in §2, the case of modulus 3 must be excluded.

The group-matrix for  $g_{24}$  may be written

$$\begin{pmatrix} (x) & (y) & (z) & (u) & (w) & (v) \\ (z) & (x) & (y) & (w) & (v) & (u) \\ (y) & (z) & (x) & (v) & (u) & (w) \\ (u) & (w) & (v) & (x) & (y) & (z) \\ (w) & (v) & (u) & (z) & (x) & (y) \\ (v) & (u) & (w) & (y) & (z) & (x) \end{pmatrix}, \quad (6)$$

where  $(x)$ ,  $(y)$ ,  $(z)$  are defined by (14), and

$$(u) \equiv \begin{pmatrix} u_1 & u_2 & u_4 & u_3 \\ u_2 & u_1 & u_3 & u_4 \\ u_3 & u_4 & u_1 & u_2 \\ u_4 & u_3 & u_2 & u_1 \end{pmatrix}, \quad (v) \equiv \begin{pmatrix} v_1 & v_2 & v_4 & v_3 \\ v_4 & v_3 & v_1 & v_2 \\ v_2 & v_1 & v_3 & v_4 \\ v_3 & v_4 & v_2 & v_1 \end{pmatrix}, \quad (w) \equiv \begin{pmatrix} w_1 & w_4 & w_3 & w_2 \\ w_2 & w_3 & w_4 & w_1 \\ w_3 & w_2 & w_1 & w_4 \\ w_4 & w_1 & w_2 & w_3 \end{pmatrix}.$$

Consider the transformation  $T$  on  $\xi_1, \dots, \xi_{24}$  with the matrix of coefficients (6). If the field  $F$  does not have modulus 2, we introduce as new variables  $\eta_1, \dots, \eta_{12}$ , defined as in § 5, and  $\eta_{13}, \dots, \eta_{24}$  defined so that  $\eta_{12+j}$  is derived from  $\eta_j$  upon adding 12 to the subscripts of  $\xi_1, \dots, \xi_{12}$ . Thus  $\eta_{13} = \xi_{13} + \xi_{14} - \xi_{15} - \xi_{16}$ . Then  $T$  gives rise to a transformation on  $\eta_1, \eta_2, \eta_3, \eta_{13}, \eta_{14}, \eta_{15}$  with the matrix of coefficients  $C_1$ :

$$C_1 \equiv \begin{pmatrix} X_1 & Y_1 & Z_1 & U_1 & W_1 & V_1 \\ Z_2 & X_2 & Y_2 & W_2 & V_2 & U_2 \\ Y_3 & Z_3 & X_3 & V_3 & U_3 & W_3 \\ U_1 & W_1 & V_1 & X_1 & Y_1 & Z_1 \\ W_2 & V_2 & U_2 & Z_2 & X_2 & Y_2 \\ V_3 & U_3 & W_3 & Y_3 & Z_3 & X_3 \end{pmatrix}, \quad C_4 \equiv \begin{pmatrix} X_4 & Y_4 & Z_4 & U_4 & W_4 & V_4 \\ Z_4 & X_4 & Y_4 & W_4 & V_4 & U_4 \\ Y_4 & Z_4 & X_4 & V_4 & U_4 & W_4 \\ U_4 & W_4 & V_4 & X_4 & Y_4 & Z_4 \\ W_4 & V_4 & U_4 & Z_4 & X_4 & Y_4 \\ V_4 & U_4 & W_4 & Y_4 & Z_4 & X_4 \end{pmatrix}.$$

Also,  $T$  gives rise to a transformation on  $\eta_4, \eta_5, \eta_6, \eta_{19}, \eta_{20}, \eta_{21}$  with the matrix of coefficients  $C_1$ ; a transformation on  $\eta_7, \eta_8, \eta_9, \eta_{16}, \eta_{17}, \eta_{18}$  with matrix  $C_1$ ; a transformation on  $\eta_{10}, \eta_{11}, \eta_{12}, \eta_{22}, \eta_{23}, \eta_{24}$  with matrix  $C_4$ . We next set

$$\begin{aligned} \zeta_1 &= \eta_1 + \eta_{13}, & \zeta_2 &= \eta_2 + \eta_{14}, & \zeta_3 &= \eta_3 + \eta_{15}, & \zeta_{13} &= \eta_1 - \eta_{13}, & \zeta_{14} &= \eta_2 - \eta_{14}, & \zeta_{15} &= \eta_3 - \eta_{15}, \\ \zeta_4 &= \eta_4 + \eta_{19}, & \zeta_5 &= \eta_5 + \eta_{20}, & \zeta_6 &= \eta_6 + \eta_{21}, & \zeta_{19} &= \eta_4 - \eta_{19}, & \zeta_{20} &= \eta_5 - \eta_{20}, & \zeta_{21} &= \eta_6 - \eta_{21}, \\ \zeta_7 &= \eta_7 + \eta_{16}, & \zeta_8 &= \eta_8 + \eta_{17}, & \zeta_9 &= \eta_9 + \eta_{18}, & \zeta_{16} &= \eta_7 - \eta_{16}, & \zeta_{17} &= \eta_8 - \eta_{17}, & \zeta_{18} &= \eta_9 - \eta_{18}. \end{aligned}$$

Then  $T$  gives rise to a transformation on  $\zeta_1, \zeta_2, \zeta_3$  with the matrix  $P$ :

$$P \equiv \begin{pmatrix} X_1 + U_1 & Y_1 + W_1 & Z_1 + V_1 \\ Z_2 + W_2 & X_2 + V_2 & Y_2 + U_2 \\ Y_3 + V_3 & Z_3 + U_3 & X_3 + W_3 \end{pmatrix}, \quad Q \equiv \begin{pmatrix} X_1 - U_1 & Y_1 - W_1 & Z_1 - V_1 \\ Z_2 - W_2 & X_2 - V_2 & Y_2 - U_2 \\ Y_3 - V_3 & Z_3 - U_3 & X_3 - W_3 \end{pmatrix}.$$

Also,  $T$  gives rise to a transformation on  $\zeta_4, \zeta_5, \zeta_6$  with matrix  $P$ ; a transformation on  $\zeta_7, \zeta_8, \zeta_9$  with matrix  $P$ ; a transformation on  $\zeta_{13}, \zeta_{14}, \zeta_{15}$  with matrix  $Q$ ; a transformation on  $\zeta_{16}, \zeta_{17}, \zeta_{18}$  with matrix  $Q$ ; a transformation on  $\zeta_{19}, \zeta_{20}, \zeta_{21}$  with matrix  $Q$ .

Applying to the variables  $\eta_{10}, \eta_{11}, \eta_{12}, \eta_{22}, \eta_{23}, \eta_{24}$  a transformation whose matrix of coefficients is the matrix  $B$  of § 3, matrix  $C_4$  is transformed into matrix  $C$  of § 3, viz.:

$$C \equiv \begin{bmatrix} R & O & O \\ O & S & O \\ O & O & S \end{bmatrix}, \quad R \equiv \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}, \quad S \equiv \begin{bmatrix} X_4 - Y_4 + V_4 - W_4 & Y_4 - Z_4 + U_4 - V_4 \\ Z_4 - Y_4 + U_4 - W_4 & X_4 - Z_4 + W_4 - V_4 \end{bmatrix},$$

$$\lambda \equiv X_4 + Y_4 + Z_4 + U_4 + W_4 + V_4, \quad \mu \equiv X_4 + Y_4 + Z_4 - U_4 - W_4 - V_4.$$

Within a field  $F$  not having modulus 2 or 3, the group-matrix (6) of the symmetric group on four letters may be transformed into a canonical matrix which may be exhibited compactly as an aggregate of  $10^2$  matrices, the 10 in the main diagonal being  $P, P, P, Q, Q, Q, \lambda, \mu, S, S$ , while the remaining 90 matrices have all their elements zero. Here  $P$  and  $Q$  are square matrices of order 3,  $S$  is a square matrix of order 2,  $\lambda$  and  $\mu$  are elements (matrices of order 1). The 9 functions giving the elements of  $P$ , the 9 giving the elements of  $Q$ , the 4 giving the elements of  $S$ , and the 2 functions  $\lambda, \mu$  furnish 24 linearly independent functions of  $x_i, y_i, z_i, u_i, v_i, w_i$ . The determinants of  $P, Q, S$  are irreducible functions.

#### THE ALTERNATING GROUP ON FIVE LETTERS

7. This group  $g_{60}$  may be generated by the substitutions:\*

$$\alpha_1 = (12345), \quad \beta_0 = (14)(23), \quad \gamma_{00} = (12)(34).$$

The following notations will be employed:

$$\begin{aligned} \alpha_0 &= \text{identity}, \quad \alpha_i = \alpha_i^i, & \gamma_{j0} &= \alpha_j^{-1} \gamma_{00}, & \gamma_{ji} &= \gamma_{j0} \alpha_i, \\ \beta_i &= \beta_0 \alpha_i, & \delta_{00} &= \gamma_{00} \beta_0, & \delta_{j0} &= \alpha_j^{-1} \delta_{00}, & \delta_{ji} &= \delta_{j0} \alpha_i. \end{aligned}$$

The sixty substitutions of  $g_{60}$  are given by the formulæ:

$$\alpha_i, \beta_i, \gamma_{ji}, \delta_{ji} \quad (i, j = 0, 1, 2, 3, 4). \quad (7)$$

They fall into five distinct sets of conjugates within  $g_{60}$ :

$$\begin{aligned} &\{\alpha_0\}; \\ &\{\gamma_{01} = (135), \gamma_{10} = (153), \gamma_{04} = (254), \gamma_{40} = (245), \gamma_{12} = (124), \gamma_{21} = (142), \\ &\quad \gamma_{23} = (235), \gamma_{32} = (253), \gamma_{34} = (134), \gamma_{43} = (143), \delta_{02} = (152), \delta_{20} = (125), \delta_{03} = (345), \\ &\quad \delta_{13} = (132), \delta_{31} = (123), \delta_{14} = (145), \delta_{41} = (154), \delta_{24} = (243), \delta_{42} = (234), \delta_{30} = (354)\}; \\ &\{\beta_0, \beta_1 = (15)(24), \beta_2 = (25)(34), \beta_3 = (12)(35), \beta_4 = (13)(45), \\ &\quad \gamma_{00} = (12)(34), \gamma_{11} = (23)(45), \gamma_{22} = (15)(34), \gamma_{33} = (12)(45), \gamma_{44} = (15)(23), \\ &\quad \delta_{00} = (13)(24), \delta_{11} = (24)(35), \delta_{22} = (14)(35), \delta_{33} = (14)(25), \delta_{44} = (13)(25)\}; \\ &\{ \alpha_1, \gamma_{02} = (14523), \gamma_{03} = (15324), \gamma_{13} = (13425), \gamma_{14} = (14352), \gamma_{24} = (12453), \} \\ &\{ \alpha_4, \gamma_{30} = (13254), \gamma_{30} = (14235), \gamma_{31} = (15243), \gamma_{41} = (12534), \gamma_{43} = (13542) \}; \\ &\{ \alpha_2, \delta_{01} = (14325), \delta_{04} = (12354), \delta_{12} = (12543), \delta_{23} = (15423), \delta_{44} = (15342), \} \\ &\{ \alpha_3, \delta_{10} = (15234), \delta_{40} = (14532), \delta_{31} = (13452), \delta_{23} = (13245), \delta_{43} = (12435) \}. \end{aligned}$$

The group-matrix of  $g_{60}$  may be written in the compact form (4), with the amplification (5), where  $x_i, y_i, z_i$  themselves now denote matrices, as follows:

\* They may be identified with  $S, U, T$ , respectively, of KLEIN, *Ikosæder*, p. 26.

$$\begin{aligned}
x_1 &= \begin{pmatrix} a_0 & a_1 & a_2 & a_3 & a_4 \\ a_4 & a_0 & a_1 & a_2 & a_3 \\ a_3 & a_4 & a_0 & a_1 & a_2 \\ a_2 & a_3 & a_4 & a_0 & a_1 \\ a_1 & a_2 & a_3 & a_4 & a_0 \end{pmatrix}, & x_4 &= \begin{pmatrix} \beta_0 & \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ \beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_0 \\ \beta_2 & \beta_3 & \beta_4 & \beta_0 & \beta_1 \\ \beta_3 & \beta_4 & \beta_0 & \beta_1 & \beta_2 \\ \beta_4 & \beta_0 & \beta_1 & \beta_2 & \beta_3 \end{pmatrix}, \\
x_2 &= \begin{pmatrix} \gamma_{00} & \gamma_{01} & \gamma_{02} & \gamma_{03} & \gamma_{04} \\ \gamma_{10} & \gamma_{11} & \gamma_{12} & \gamma_{13} & \gamma_{14} \\ \gamma_{20} & \gamma_{21} & \gamma_{22} & \gamma_{23} & \gamma_{24} \\ \gamma_{30} & \gamma_{31} & \gamma_{32} & \gamma_{33} & \gamma_{34} \\ \gamma_{40} & \gamma_{41} & \gamma_{42} & \gamma_{43} & \gamma_{44} \end{pmatrix}, & x_3 &= \begin{pmatrix} \delta_{00} & \delta_{01} & \delta_{02} & \delta_{03} & \delta_{04} \\ \delta_{10} & \delta_{11} & \delta_{12} & \delta_{13} & \delta_{14} \\ \delta_{20} & \delta_{21} & \delta_{22} & \delta_{23} & \delta_{24} \\ \delta_{30} & \delta_{31} & \delta_{32} & \delta_{33} & \delta_{34} \\ \delta_{40} & \delta_{41} & \delta_{42} & \delta_{43} & \delta_{44} \end{pmatrix}, \\
y_1 &= (\delta_{31}), & y_2 &= (\delta_{34}), & y_3 &= (\gamma_{31}), & y_4 &= (\gamma_{34}), \\
z_1 &= (\delta_{13}), & z_2 &= (\delta_{43}), & z_3 &= (\gamma_{13}), & z_4 &= (\gamma_{43}).
\end{aligned}$$

Each of the eight last matrices is designated by the first element in its first row, since the remaining elements are then determined by a cyclic permutation of the subscripts 0, 1, 2, 3, 4. For example, the first row of  $y_1$  is  $\delta_{31}, \delta_{32}, \delta_{33}, \delta_{34}, \delta_{30}$ ; the second row is  $\delta_{41}, \delta_{42}, \delta_{43}, \delta_{44}, \delta_{40}$ ; the third row is  $\delta_{01}, \delta_{02}, \delta_{03}, \delta_{04}, \delta_{00}$ ; etc. Note that, if we replace each matrix by the first element in its first row, we obtain the group-matrix of the subgroup  $g_{12}$  of the even substitutions on the letters 1, 2, 3, 4.

8. Consider a transformation  $T$  on the variables  $\xi_0, \xi_1, \dots, \xi_{39}$ , the matrix of whose coefficients is the group-matrix of  $g_{80}$ . Let  $\omega$  be a primitive fifth root of unity and introduce the new variables  $\eta_{ji}$  ( $j=1, \dots, 12; i=0, 1, 2, 3, 4$ ), where

$$\eta_{1i} = \xi_0 + \omega^i \xi_1 + \omega^{2i} \xi_2 + \omega^{3i} \xi_3 + \omega^{4i} \xi_4, \quad \eta_{7i} = \xi_{15} + \omega^i \xi_{16} + \omega^{2i} \xi_{17} + \omega^{3i} \xi_{18} + \omega^{4i} \xi_{19},$$

the expressions for the remaining  $\eta_{ji}$  being given in the compact form:

$$\begin{array}{c}
\begin{array}{ccccc}
1 & \omega^{2i} & \omega^{3i} & \omega^{4i} & \omega^i \\
\hline
\eta_{2i} & \xi_5 & \xi_{34} & \xi_{36} & \xi_{47} & \xi_{58} \\
\eta_{3i} & \xi_6 & \xi_{30} & \xi_{27} & \xi_{48} & \xi_{59} \\
\eta_{4i} & \xi_7 & \xi_{31} & \xi_{28} & \xi_{49} & \xi_{50} \\
\eta_{5i} & \xi_8 & \xi_{32} & \xi_{29} & \xi_{50} & \xi_{56} \\
\eta_{6i} & \xi_9 & \xi_{33} & \xi_{35} & \xi_{46} & \xi_{57}
\end{array}
, &
\begin{array}{ccccc}
1 & \omega^{3i} & \omega^{2i} & \omega^i & \omega^{4i} \\
\hline
\eta_{8i} & \xi_{10} & \xi_{24} & \xi_{26} & \xi_{43} & \xi_{53} \\
\eta_{9i} & \xi_{11} & \xi_{20} & \xi_{27} & \xi_{43} & \xi_{54} \\
\eta_{10i} & \xi_{12} & \xi_{21} & \xi_{28} & \xi_{44} & \xi_{50} \\
\eta_{11i} & \xi_{13} & \xi_{22} & \xi_{29} & \xi_{40} & \xi_{51} \\
\eta_{12i} & \xi_{14} & \xi_{23} & \xi_{25} & \xi_{41} & \xi_{52}
\end{array}
\end{array}$$

Then  $T$  may be written as the product  $T_0 T_1 T_2 T_3 T_4$ , where  $T_i$  is a transformation on the twelve variables  $\eta_{1i}, \dots, \eta_{12i}$ , with the matrix of coefficients

$$(T_i) = \begin{pmatrix} \phi_i & \psi_i \\ \psi_{-i} & \phi_{-i} \end{pmatrix},$$

where  $\phi_i$  and  $\psi_i$  denote the respective matrices

$$\begin{pmatrix} A_i & C_{0i} & C_{1i} & C_{2i} & C_{3i} & C_{4i} \\ \Gamma_{0i} & F_{0i} & F_{1i} & F_{2i} & F_{3i} & F_{4i} \\ \Gamma_{1i} & K_{4i} & K_{0i} & K_{1i} & K_{2i} & K_{3i} \\ \Gamma_{2i} & R_{3i} & R_{4i} & R_{0i} & R_{1i} & R_{2i} \\ \Gamma_{3i} & M_{2i} & M_{3i} & M_{4i} & M_{0i} & M_{1i} \\ \Gamma_{4i} & E_{1i} & E_{2i} & E_{3i} & E_{4i} & E_{0i} \end{pmatrix}, \quad \begin{pmatrix} B_i & D_{0i} & D_{1i} & D_{2i} & D_{3i} & D_{4i} \\ \Delta_{0i} & G_{0i} & G_{1i} & G_{2i} & G_{3i} & G_{4i} \\ \Delta_{1i} & J_{1i} & J_{2i} & J_{3i} & J_{4i} & J_{0i} \\ \Delta_{2i} & S_{2i} & S_{3i} & S_{4i} & S_{0i} & S_{1i} \\ \Delta_{3i} & N_{3i} & N_{4i} & N_{0i} & N_{1i} & N_{2i} \\ \Delta_{4i} & L_{4i} & L_{0i} & L_{1i} & L_{2i} & L_{3i} \end{pmatrix},$$

in which the following abbreviations are used:

$$\begin{aligned}
 A_i &= \sum_{k=0}^4 \omega^{ki} a_k, & B_i &= \sum_{k=0}^4 \omega^{-ki} \beta_k, & C_{ji} &= \sum_{k=0}^4 \omega^{-ki} \gamma_{kj}, \\
 D_{ji} &= \sum_{k=0}^4 \omega^{-ki} \delta_{kj}, & \Delta_{ji} &= \sum_{k=0}^4 \omega^{-ki} \delta_{jk}, & \Gamma_{ji} &= \sum_{k=0}^4 \omega^{ki} \gamma_{jk},
 \end{aligned}$$

	1	$\omega^i$	$\omega^{4i}$	$\omega^{3i}$	$\omega^{2i}$		1	$\omega^i$	$\omega^{4i}$	$\omega^{3i}$	$\omega^{2i}$
$F_{0i}$	$\alpha_0$	$\gamma_{14}$	$\gamma_{41}$	$\delta_{32}$	$\delta_{23}$	$G_{0i}$	$\beta_0$	$\delta_{11}$	$\delta_{44}$	$\gamma_{33}$	$\gamma_{22}$
$K_{0i}$	$\alpha_0$	$\gamma_{30}$	$\gamma_{03}$	$\delta_{43}$	$\delta_{34}$	$J_{0i}$	$\beta_0$	$\delta_{30}$	$\delta_{03}$	$\gamma_{42}$	$\gamma_{31}$
$R_{0i}$	$\alpha_0$	$\gamma_{31}$	$\gamma_{13}$	$\delta_{04}$	$\delta_{40}$	$S_{0i}$	$\beta_0$	$\delta_{34}$	$\delta_{12}$	$\gamma_{01}$	$\gamma_{40}$
$M_{0i}$	$\alpha_0$	$\gamma_{43}$	$\gamma_{24}$	$\delta_{10}$	$\delta_{01}$	$N_{0i}$	$\beta_0$	$\delta_{43}$	$\delta_{21}$	$\gamma_{10}$	$\gamma_{04}$
$E_{0i}$	$\alpha_0$	$\gamma_{03}$	$\gamma_{30}$	$\delta_{21}$	$\delta_{12}$	$L_{0i}$	$\beta_0$	$\delta_{02}$	$\delta_{20}$	$\gamma_{24}$	$\gamma_{13}$

From  $F_{0i}$  we obtain  $F_{1i}$  by adding unity to the subscript of  $\alpha$  and to the second subscripts of  $\gamma, \delta$ ;  $F_{2i}$  by adding 2 to the same subscripts; etc. Thus

$$F_{1i} = \alpha_1 + \omega^i \gamma_{10} + \omega^{4i} \gamma_{43} + \omega^{3i} \delta_{33} + \omega^{2i} \delta_{24}.$$

By a similar process, we obtain  $K_{ji}, R_{ji}, M_{ji}, E_{ji}, G_{ji}, \dots, L_{ji}$ .

Now  $T_i$  is transformed into  $T_{-i} \equiv T_{5-i}$  by the substitution

$$(\eta_{1i} \eta_{i, -i}) (\eta_{2i} \eta_{i, -i}) (\eta_{3i} \eta_{i, -i}) (\eta_{4i} \eta_{i, -i}) (\eta_{5i} \eta_{i, -i}) \dots (\eta_{6i} \eta_{i, -i}) (\eta_{12i} \eta_{i, -i}).$$

Hence  $T$  may be transformed into the product  $T_0 T_1 \bar{T}_1 T_2 \bar{T}_2$ , where  $\bar{T}_j$  has the same matrix as  $T_j$ , but affects different variables. Now

$$(T_0) = \begin{pmatrix} \phi_0 & \psi_0 \\ \psi_0 & \phi_0 \end{pmatrix}.$$

Hence, if we introduce in place of  $\eta_{j0}$  ( $j = 1, \dots, 12$ ) the variables

$$X_j = \eta_{j0} + \eta_{6+j, 0}, \quad Y_j = \eta_{j0} - \eta_{6+j, 0} \quad (j = 1, \dots, 6),$$

we obtain for  $T_0$  the product  $XY$ , where  $X$  is the transformation on  $X_1, \dots, X_6$ , whose matrix of coefficients is  $M_+$ , and  $Y$  is the transformation on  $Y_1, \dots, Y_6$ , whose matrix of coefficients is  $M_-$ ,  $M_{\pm}$  denoting the following matrix:

$$\begin{pmatrix}
 A_0 \pm B_0 & C_{00} \pm D_{00} & C_{10} \pm D_{10} & C_{20} \pm D_{20} & C_{30} \pm D_{30} & C_{40} \pm D_{40} \\
 \Gamma_{00} \pm \Delta_{00} & F_{00} \pm G_{00} & F_{10} \pm G_{10} & F_{20} \pm G_{20} & F_{30} \pm G_{30} & F_{40} \pm G_{40} \\
 \Gamma_{10} \pm \Delta_{10} & K_{40} \pm J_{10} & K_{00} \pm J_{30} & K_{10} \pm J_{20} & K_{20} \pm J_{40} & K_{30} \pm J_{00} \\
 \Gamma_{20} \pm \Delta_{20} & R_{30} \pm S_{20} & R_{40} \pm S_{30} & R_{00} \pm S_{40} & R_{10} \pm S_{00} & R_{20} \pm S_{10} \\
 \Gamma_{30} \pm \Delta_{30} & M_{20} \pm N_{30} & M_{30} \pm N_{40} & M_{40} \pm N_{00} & M_{00} \pm N_{10} & M_{10} \pm N_{20} \\
 \Gamma_{40} \pm \Delta_{40} & E_{10} \pm L_{40} & E_{20} \pm L_{00} & E_{30} \pm L_{10} & E_{40} \pm L_{20} & E_{00} \pm L_{30}
 \end{pmatrix}.$$

9. Further reductions are found most simply by considering the special group-matrix  $T'$ , obtained by equating elements which correspond to conjugate operators of  $g_{60}$ . We therefore set

$$a = a_0, \quad c = \gamma_{01} = \gamma_{10} = \dots, \quad b = \beta_0 = \beta_1 = \dots, \quad d = a_1 = a_4 = \dots, \quad e = a_2 = a_3 = \dots.$$

We obtain the results:

$$A_i = a + d(\omega^i + \omega^{4i}) + e(\omega^{2i} + \omega^{3i}), \quad B_0 = 5b, \quad B_1 = B_2 = B_3 = B_4 = 0,$$

$$\begin{array}{c} \begin{array}{cc} & \begin{array}{ccccc} 1 & \omega^{-1} & \omega^{-2i} & \omega^{-3i} & \omega^{-4i} \end{array} \\ \begin{array}{c} D_{0i}, \Delta_{0i} \\ D_{1i}, \Delta_{1i} \\ D_{2i}, \Delta_{2i} \\ D_{3i}, \Delta_{3i} \\ D_{4i}, \Delta_{4i} \end{array} & \begin{array}{ccccc} \left| \begin{array}{ccccc} b & e & c & c & e \\ e & b & e & c & c \\ c & e & b & e & c \\ c & c & e & b & e \\ e & c & c & e & b \end{array} \right| \end{array} \end{array}, \quad \begin{array}{c} \begin{array}{cc} & \begin{array}{ccccc} 1 & \omega^{-1} & \omega^{-2i} & \omega^{-3i} & \omega^{-4i} \end{array} \\ \begin{array}{c} C_{0i}, \Gamma_{0,-i} \\ C_{1i}, \Gamma_{1,-i} \\ C_{2i}, \Gamma_{2,-i} \\ C_{3i}, \Gamma_{3,-i} \\ C_{4i}, \Gamma_{4,-i} \end{array} & \begin{array}{ccccc} \left| \begin{array}{ccccc} b & c & d & d & c \\ c & b & c & d & d \\ d & c & b & c & d \\ d & d & c & b & c \\ c & d & d & c & b \end{array} \right| \end{array} \end{array}, \end{array}$$

$$\begin{array}{c} \begin{array}{cc} & \begin{array}{ccccc} 1 & \omega^i & \omega^{4i} & \omega^{2i} & \omega^{3i} \end{array} \\ \begin{array}{c} F_{0i}, K_{0i}, R_{0i}, M_{0i}, E_{0i} \\ F_{1i}, K_{1i}, R_{1i}, M_{1i}, E_{1i} \\ F_{2i}, K_{2i}, R_{2i}, M_{2i}, E_{2i} \\ F_{3i}, K_{3i}, R_{3i}, M_{3i}, E_{3i} \\ F_{4i}, K_{4i}, R_{4i}, M_{4i}, E_{4i} \end{array} & \begin{array}{ccccc} \left| \begin{array}{ccccc} a & d & d & e & e \\ d & c & d & b & c \\ e & b & c & e & c \\ e & c & b & c & e \\ d & d & c & c & b \end{array} \right| \end{array} \end{array}, \end{array}$$

$$\begin{array}{c} \begin{array}{cc} & \begin{array}{ccccc} 1 & \omega^i & \omega^{4i} & \omega^{2i} & \omega^{3i} \end{array} \\ \begin{array}{c} G_{0i}, J_{2i}, S_{4i}, N_{1i}, L_{3i} \\ G_{1i}, J_{3i}, S_{0i}, N_{2i}, L_{4i} \\ G_{2i}, J_{4i}, S_{1i}, N_{3i}, L_{0i} \\ G_{3i}, J_{0i}, S_{2i}, N_{4i}, L_{1i} \\ G_{4i}, J_{1i}, S_{3i}, N_{0i}, L_{2i} \end{array} & \begin{array}{ccccc} \left| \begin{array}{ccccc} b & b & b & b & b \\ b & e & e & c & c \\ b & c & c & d & d \\ b & c & c & d & d \\ b & e & e & c & c \end{array} \right| \end{array} \end{array}.$$

10. The transformation  $X'$  with the matrix  $M'_+$  takes the form

$$X'_j = xX_j + y \sum_{\substack{k=1, \dots, 6 \\ k \neq j}} X_k \quad (j=1, \dots, 6),$$

where we have set

$$x = a + 5b + 2d + 2e, \quad y = 2b + 4c + 2d + 2e.$$

Introduce new variables by the transformation of determinant 6

$$\zeta_1 = X_1 + X_2 + X_3 + X_4 + X_5 + X_6, \quad \zeta_j = X_j - X_1 \quad (j=2, \dots, 6).$$

Then  $X'$  multiplies  $\zeta_1$  by  $r_0$ , and each  $\zeta_j$  by  $r_1$ , where

$$r_0 = x + 5y = a + 15b + 20c + 12d + 12e, \quad r_1 = x - y = a + 3b - 4c.$$

Since  $X$ , whose matrix is  $M_+$ , is commutative with  $X'$ , it must multiply  $\zeta_1$  by a constant  $f$  and give rise to a transformation on  $\zeta_2, \dots, \zeta_6$ . The matrix of the latter is

$$\begin{pmatrix} f_0 & f_1 & f_2 & f_3 & f_4 \\ k_0 & k_1 & k_2 & k_3 & k_4 \\ r_0 & r_1 & r_2 & r_3 & r_4 \\ m_0 & m_1 & m_2 & m_3 & m_4 \\ e_0 & e_1 & e_2 & e_3 & e_4 \end{pmatrix}, \quad \begin{cases} f_j = F_{j0} + G_{j0} - C_{j0} - D_{j0} \\ k_i = K_{i0} + J_{i0} - C_{i0} - D_{i0} \\ m_2 = M_{20} + N_{20} - C_{20} - D_{20} \\ r_3 = R_{30} + S_{30} - C_{30} - D_{30} \\ e_1 = E_{10} + L_{10} - C_{10} - D_{10} \end{cases},$$

the values of  $k_j, r_j, m_j, l_j$  being derived from those of  $k_i, r_i, m_i, e_i$  by cyclic permutations of the first subscripts 0, 1, 2, 3, 4 of  $K, J, C, D, R, S, \dots$ . Finally,  $f$  is the sum of the 60 elements (7) of the initial group-matrix, and  $\zeta_1$  is the sum of the 60 initial variables  $\xi_0, \dots, \xi_{59}$ .

11. The matrix  $M'$  of the transformation  $Y'$  on  $Y_1, \dots, Y_6$  takes the form

$$\begin{pmatrix} z & w & w & w & w & w \\ w & z & w & -w & -w & w \\ w & w & z & w & -w & -w \\ w & -w & w & z & w & -w \\ w & -w & -w & w & z & w \\ w & w & -w & -w & w & z \end{pmatrix},$$

$$z \equiv a - 5b + 2d + 2e, \quad w \equiv 2d - 2e.$$

To determine a function  $a_1 Y_1 + \dots + a_6 Y_6$  which  $Y'$  multiplies by a constant  $r_4$ , we first verify that  $z - \sqrt{5}w$  is a root of the characteristic equation of  $Y'$ . Taking  $r_4 = z - \sqrt{5}w$ , we have the following conditions on  $a_1, \dots, a_6$ , after dropping the common factor  $w$ :

$$\begin{aligned} \sqrt{5}a_1 + a_2 + a_3 + a_4 + a_5 + a_6 &= 0, & a_1 + \sqrt{5}a_2 + a_3 - a_4 - a_5 + a_6 &= 0, \\ a_1 + a_2 + \sqrt{5}a_3 + a_4 - a_5 - a_6 &= 0, & a_1 - a_2 + a_3 + \sqrt{5}a_4 + a_5 - a_6 &= 0, \\ a_1 - a_2 - a_3 + a_4 + \sqrt{5}a_5 + a_6 &= 0, & a_1 + a_2 - a_3 - a_4 + a_5 + \sqrt{5}a_6 &= 0. \end{aligned}$$

Since the characteristic equation is satisfied, the determinant of these equations vanishes, so that the sixth equation may be dropped. By adding and subtracting the equations in the first line, we obtain

$$(\sqrt{5} + 1)(a_1 + a_3) + 2a_2 + 2a_5 = 0, \quad (\sqrt{5} - 1)(a_1 - a_3) + 2a_4 + 2a_6 = 0. \quad (8)$$

Adding and subtracting those in the second line, and eliminating  $a_5, a_6$  by (8), we get

$$\begin{aligned} (\sqrt{5} + 3)(a_1 + a_3) + (\sqrt{5} + 1)(a_2 + a_4) &= 0, \\ (\sqrt{5} - 1)(a_1 + a_3) - (\sqrt{5} - 3)(a_2 + a_4) &= 0. \end{aligned} \quad (9)$$

Since the determinant of the coefficients of  $a_1 + a_3$  and  $a_2 + a_4$  vanishes, one of these equations may be dropped. Multiplying the first equation (9) by  $\sqrt{5} - 3$ , the second equation (8) by  $\sqrt{5}$ , and adding to the first equation (8), we obtain the fifth of the original equations. Hence the independent conditions on the  $a_j$  are equations (8) and the first equation (9). For the values

$$a_1 = 1, \quad a_2 = a_4 = 0; \quad a_3 = 1, \quad a_5 = a_6 = 0; \quad a_2 = 1, \quad a_1 = a_3 = 0,$$

we obtain the respective functions

$$\pi_1 = Y_1 - Y_3 - y_1 Y_5 - y_1 Y_6, \quad \pi_2 = Y_2 - Y_4 - y_2 Y_5 + y_2 Y_6, \quad \pi_3 = Y_1 + y_2 Y_4 - y_2 Y_5 - Y_6, \quad (10)$$

where  $y_1 = \frac{1}{2}(\sqrt{5} - 1)$ ,  $y_2 = \frac{1}{2}(-\sqrt{5} - 1)$  are the roots of  $y^2 + y - 1 = 0$ . Changing the sign of  $\sqrt{5}$  throughout, and hence interchanging  $y_1$  with  $y_2$ , we obtain from  $\pi_1, \pi_2, \pi_3$  three functions  $\pi_4, \pi_5, \pi_6$ , respectively, which  $Y'$  multiplies by  $z + \sqrt{5}w \equiv r_3$ . Hence the canonical form of  $Y'$  is

$$\pi'_1 = r_4 \pi_1, \quad \pi'_2 = r_4 \pi_2, \quad \pi'_3 = r_4 \pi_3, \quad \pi'_4 = r_3 \pi_4, \quad \pi'_5 = r_3 \pi_5, \quad \pi'_6 = r_3 \pi_6. \quad (11)$$

That the variables  $\pi_i$  are independent follows from the solution of (10) and (10'):

$$\begin{aligned}
-2\sqrt{5}Y_1 &= -\sqrt{5}\pi_1 + \pi_2 - \sqrt{5}\pi_4 - \pi_5 - 2y_1\pi_3 + 2y_2\pi_6, & -2\sqrt{5}Y_5 &= \pi_1 - \pi_2 - \pi_4 + \pi_5, \\
-2\sqrt{5}Y_2 &= \pi_1 + \pi_2 - \pi_4 - \pi_5 - 2y_1\pi_3 + 2y_2\pi_6, & -2\sqrt{5}Y_6 &= \pi_1 + \pi_2 - \pi_4 - \pi_5, \\
-2\sqrt{5}Y_3 &= \pi_1 - \sqrt{5}\pi_2 + 2\pi_3 - \pi_4 - \sqrt{5}\pi_5 - 2\pi_6, & & \\
& & -2\sqrt{5}Y_4 &= \pi_1 - \pi_2 + 2\pi_3 - \pi_4 + \pi_5 - 2\pi_6.
\end{aligned}$$

The values of  $r_3$  and  $r_4$  are as follows:

$$r_3 = z + \sqrt{5}w = a - 5b - 4y_1d - 4y_2e, \quad r_4 = z - \sqrt{5}w = a - 5b - 4y_1d - 4y_2e.$$

Since the transformation  $Y$  of matrix  $M_-$  is commutative with  $Y'$  of matrix  $M'_-$ ,  $Y$  must equal the product of a transformation  $C$  on  $\pi_1, \pi_2, \pi_3$ , and a transformation  $\bar{C}$  on  $\pi_4, \pi_5, \pi_6$ , the matrix of  $\bar{C}$  being derived from the matrix of  $C$  by changing the sign of  $\sqrt{5}$ .

12. For the special group-matrix,  $\phi_i$  and  $\psi_i$  become respectively

$$\begin{aligned}
\phi'_i &\equiv \begin{pmatrix} F_{0i} & C_{0i} & \omega^{-1}C_{0i} & \omega^{-2i}C_{0i} & \omega^{-3i}C_{0i} & \omega^{-4i}C_{0i} \\ C_{0i} & F_{0i} & F_{1i} & F_{2i} & F_{3i} & F_{4i} \\ \omega^1 C_{0i} & F_{4i} & F_{0i} & F_{1i} & F_{2i} & F_{3i} \\ \omega^{2i} C_{0i} & F_{3i} & F_{4i} & F_{0i} & F_{1i} & F_{2i} \\ \omega^{3i} C_{0i} & F_{2i} & F_{3i} & F_{4i} & F_{0i} & F_{1i} \\ \omega^{4i} C_{0i} & F_{1i} & F_{2i} & F_{3i} & F_{4i} & F_{0i} \end{pmatrix}, \\
\psi'_i &\equiv \begin{pmatrix} 0 & D_{0i} & \omega^{-1}D_{0i} & \omega^{-2i}D_{0i} & \omega^{-3i}D_{0i} & \omega^{-4i}D_{0i} \\ D_{0i} & 0 & G_{1i} & G_{2i} & G_{3i} & G_{4i} \\ \omega^{-1}D_{0i} & G_{1i} & 0 & G_{1i} & G_{2i} & G_{3i} \\ \omega^{-2i}D_{0i} & G_{2i} & G_{1i} & 0 & G_{1i} & G_{2i} \\ \omega^{-3i}D_{0i} & G_{3i} & G_{2i} & G_{1i} & 0 & G_{1i} \\ \omega^{-4i}D_{0i} & G_{4i} & G_{3i} & G_{2i} & G_{1i} & 0 \end{pmatrix}.
\end{aligned}$$

We introduce in  $T'_1$  the new variables

$$\begin{aligned}
Z_1 &\equiv \eta_{11}, & Z_{2+j} &\equiv \eta_{21} + \omega^{-j}\eta_{31} + \omega^{-2j}\eta_{41} + \omega^{-3j}\eta_{51} + \omega^{-4j}\eta_{61} \\
Z_7 &\equiv \eta_{71}, & Z_{8+j} &\equiv \eta_{81} + \omega^{-j}\eta_{91} + \omega^{-2j}\eta_{101} + \omega^{-3j}\eta_{111} + \omega^{-4j}\eta_{121}
\end{aligned} \quad (j=0, 1, 2, 3, 4),$$

and make use of the following abbreviations:

$$\rho_j = F_{0i} + \omega^j F_{1i} + \omega^{2j} F_{2i} + \omega^{3j} F_{3i} + \omega^{4j} F_{4i}, \quad \sigma_j = \omega^j G_{1i} + \omega^{2j} G_{2i} + \omega^{3j} G_{3i} + \omega^{4j} G_{4i}.$$

Noting that the minors of the first elements of the first rows of  $\phi'_i$  and  $\psi'_i$  are cyclic matrices, we may make use of the result of § 2. Also,

$$F_{i,-i} = F_{1,i}, \quad F_{2,-i} = F_{2,i}, \quad F_{0,-i} = F_{0,i}, \quad G_{j,-i} = G_{j,i}, \quad D_{0,-i} = D_{0,i}, \quad C_{0,-i} = C_{0,i}.$$

We obtain for  $T'_1$  the following form:

$$\left. \begin{aligned}
Z'_1 &= F_{0i}Z_1 + C_{0i}Z_2 + D_{0i}Z_3, & Z'_2 &= \rho_0Z_2 + \sigma_0Z_8, & Z'_3 &= 5C_{0i}Z_1 + \rho_1Z_2 + \sigma_1Z_9, \\
Z'_4 &= \rho_2Z_4 + \sigma_2Z_{10}, & Z'_5 &= \rho_3Z_5 + \sigma_3Z_{11}, & Z'_6 &= \rho_4Z_6 + 5D_{0i}Z_7 + \sigma_4Z_{12}, \\
Z'_7 &= D_{0i}Z_8 + F_{0i}Z_9 + C_{0i}Z_{12}, & Z'_8 &= \sigma_0Z_2 + \rho_0Z_8, & Z'_9 &= 5D_{0i}Z_1 + \sigma_1Z_2 + \rho_1Z_9, \\
Z'_{10} &= \sigma_2Z_4 + \rho_2Z_{10}, & Z'_{11} &= \sigma_3Z_5 + \rho_3Z_{11}, & Z'_{12} &= \sigma_4Z_6 + 5C_{0i}Z_7 + \rho_1Z_{12}.
\end{aligned} \right\} \quad (12)$$

Now,  $Z_2$  and  $Z_8$  occur only in  $Z'_2$  and  $Z'_8$ . Setting

$$W_2 = Z_2 + Z_8, \quad W_8 = Z_2 - Z_8, \quad (13)$$

we get

$$W'_2 = (\rho_0 + \sigma_0)W_2, \quad W'_8 = (\rho_0 - \sigma_0)W_8. \quad (14)$$



Now,  $\omega + \omega^4$  and  $\omega^2 + \omega^3$  are the roots of  $y^2 + y - 1 = 0$ . By proper choice of  $\omega$ , we may set  $\omega + \omega^4 = y_1$ ,  $\omega^2 + \omega^3 = y_2$ , where  $y_1$  and  $y_2$  are given in § 11. Then

$$\begin{aligned}\rho_0 &= a - b - 2c + 2d(1 + y_1) + 2e(1 + y_2), & \sigma_0 &= 4b - 2c + 2d y_2 + 2e y_1, \\ \rho_0 + \sigma_0 &= a + 3b - 4c = r_1, & \rho_0 - \sigma_0 &= a - 5b - 4d y_2 - 4e y_1 = r_2.\end{aligned}$$

Next,  $Z_4$  and  $Z_{10}$  occur in (12) only in  $Z'_4$  and  $Z'_{10}$ , and

$$Z'_4 = \rho_2 Z_4 + \sigma_2 Z_{10}, \quad Z'_{10} = \sigma_2 Z_4 + \rho_2 Z_{10}. \quad (15)$$

Its characteristic equation, with  $x$  as variable, is

$$x^2 - x(\rho_2 + \rho_2) + (\rho_2 \rho_2 - \sigma_2^2) = 0. \quad (16)$$

We find that

$$\begin{aligned}\rho_2 &= a + b(y_1 + 2) + c(3y_2 + 2) + d(y_1 - 1) + e(y_1 - 1), \\ \rho_2 &= a + b(y_2 + 2) + c(3y_1 + 2) + d(y_2 - 1) + e(y_2 - 1), \\ \sigma_2 &= 3c - b - d - e, \quad \rho_2 + \rho_2 = 2a + 3b + c - 3d - 3e, \\ \rho_2 \rho_2 - \sigma_2^2 &= a^2 + 3ab + ac - 3ad - 3ae + 15bc - 9bd - 9be - 20c^2 + 12cd + 12ce.\end{aligned}$$

Hence the roots of (16) are

$$r_1 = a + 3b - 4c, \quad r_2 = a + 5c - 3d - 3e.$$

Hence (15) multiplies  $\sigma_2 Z_4 - (\rho_2 - r_1) Z_{10}$  by  $r_1$ . But

$$\rho_2 - r_1 = (y_1 - 1)(b - 3c + d + e), \quad \rho_2 - r_2 = (y_1 + 2)(b - 3c + d + e),$$

so that each is a multiple of  $\sigma_2$ . Hence the transformation of variables

$$W_4 = Z_4 + (y_1 - 1) Z_{10}, \quad W_{10} = Z_4 + (y_1 + 2) Z_{10} \quad (17)$$

reduces (15) to the canonical form

$$W'_4 = r_1 W_4, \quad W'_{10} = r_2 W_{10}. \quad (18)$$

Since  $\sigma_2 = \sigma_3$ ,  $\sigma_1 = \sigma_4$ , (12) affects  $Z_5$  and  $Z_{11}$  thus:

$$Z'_{11} = \rho_3 Z_{11} + \sigma_3 Z_5, \quad Z'_5 = \sigma_3 Z_{11} + \rho_3 Z_5.$$

Hence for the new variables

$$W_{11} = Z_{11} + (y_1 - 1) Z_5, \quad W_5 = Z_{11} + (y_1 + 2) Z_5, \quad (19)$$

$$W'_{11} = r_1 W_{11}, \quad W'_5 = r_2 W_5. \quad (20)$$

Next,  $Z_1, Z_3, Z_9$  occur in (12) only in  $Z'_1, Z'_3, Z'_9$ , the matrix of coefficients being

$$\begin{pmatrix} F_{01} & C_{01} & D_{01} \\ 5C_{01} & \rho_1 & \sigma_1 \\ 5D_{01} & \sigma_1 & \rho_1 \end{pmatrix}. \quad (21)$$

But, for the special group-matrix,<sup>†</sup>

$$\begin{aligned}F_{01} &= a + d y_1 + e y_2, & C_{01} &= b + c y_1 + d y_2, & D_{01} &= b + c y_2 + e y_1, \\ \rho_1 &= a + 2b y_2 + c(y_1 + 1) + 2d(y_1 + 1) + e(y_2 - 1), \\ \rho_1 &= a + 2b y_1 + c(y_2 + 1) + d(y_1 - 1) + 2e(y_2 + 1), \\ \sigma_1 &= y_1 G_{11} + y_2 G_{21} = -b - 2c + d(1 - y_2) + e(1 - y_1).\end{aligned}$$

<sup>†</sup>As a check, it was verified that  $\rho_0 + \rho_1 + \rho_2 + \rho_3 + \rho_4 = 5F_{01}$ ,  $\sigma_0 + 2\sigma_1 + 2\sigma_2 = 0$ .

The roots of the characteristic equation of matrix (21) are seen to be  $r_1, r_2, r_3$ . Introducing the new variables

$$W_1 = Z_1 + y_1^2 Z_3 + y_2^2 Z_5, \quad W_2 = Z_1 + \sqrt{5} Z_3 - \sqrt{5} Z_5, \\ W_3 = Z_1 + (y_1 - 2) Z_3 + (y_1 - 2) Z_5,$$

the transformation with the matrix (21) becomes

$$W'_1 = r_1 W_1, \quad W'_2 = r_2 W_2, \quad W'_3 = r_3 W_3. \quad (22)$$

Finally,  $Z_7, Z_{12}, Z_6$  occur in (12) only in  $Z'_7, Z'_{12}, Z'_6$ , the matrix of coefficients in this order being (21). Hence, if we set

$$W_7 = Z_7 + y_1^2 Z_{12} + y_2^2 Z_6, \quad W_{12} = Z_7 + \sqrt{5} Z_{12} - \sqrt{5} Z_6, \\ W_6 = Z_7 + (y_1 - 2) Z_{12} + (y_1 - 2) Z_6,$$

the transformation on  $Z_7, Z_{12}, Z_6$  becomes

$$W'_7 = r_1 W_7, \quad W'_{12} = r_2 W_{12}, \quad W'_6 = r_3 W_6. \quad (23)$$

Transforming the transformation given by (14), (18), (20), (22), (23), by

$$(W_2, W_7, W_6) (W_3, W_{12}),$$

we obtain the canonical form of  $T'_1$ :

$$W'_i = r_1 W_i \quad (i=1, 2, 3, 4, 5); \quad W'_j = r_2 W_j \quad (j=6, 7, 8); \quad W'_k = r_3 W_k \quad (k=9, 10, 11, 12).$$

Now,  $T_2$  is obtained from  $T_1$  by replacing  $\omega$  by  $\omega^2$  in the variables and coefficients. Hence  $T'_2$  may be given the canonical form

$$V'_i = r_1 V_i \quad (i=1, 2, 3, 4, 5); \quad V'_j = r_2 V_j \quad (j=6, 7, 8); \quad V'_k = r_3 V_k \quad (k=9, 10, 11, 12),$$

where, as in § 11,

$$r_i \equiv a - 5b - 4d y_1 - 4e y_2.$$

Since  $T_1$  is commutative with  $T'_1$ , it is the product of three transformations on the respective sets of variables  $W_i$  ( $i=1, \dots, 5$ ),  $W_j$  ( $j=6, 7, 8$ ),  $W_k$  ( $k=9, 10, 11, 12$ ). Likewise,  $T_2$  is the product of three transformations on 5, 3, 4 variables respectively.

*For the simple group  $g_{60}$ , the special group-determinant has exactly the five distinct linear factors:*<sup>a</sup>

$$r_0 = a + 15b + 20c + 12d + 12e, \quad r_1 = a + 3b - 4c, \\ r_2 = a + 5c - 3d - 3e, \quad r_3 = a - 5b - 4y_2 d - 4y_1 e, \quad r_4 = a - 5b - 4y_1 d - 4y_2 e,$$

occurring to the respective powers 1, 25, 16, 9, 9.

The investigation, moreover, furnishes a transformation which transforms  $T$  into a product of transformations on sets of 1, 5, 4, 3 distinct variables. The general group-determinant equals  $\phi_5^5 \phi_4^4 \phi_3^3 \phi_2^2 \phi_1$ , where  $\phi_i$  is a determinant of order  $i$  with  $i^2$  arbitrary elements. The group is the direct product of a general quinary group in  $F$ , a general quaternary group in  $F$ , a general ternary group in  $F(\sqrt{5})$ , and a general unary group in  $F$ .

<sup>a</sup>This result is in accord with that of Frobenius, *Berliner Sitzungsberichte*, 1896, p. 1012.



## **ON AMORPHOUS SULPHUR**



# ON AMORPHOUS SULPHUR

ALEXANDER SMITH AND WILLIS B. HOLMES

## PART I. THE RELATION OF AMORPHOUS SULPHUR TO THE FREEZING-POINT OF LIQUID SULPHUR<sup>1</sup>

THE relation of amorphous sulphur to liquid sulphur is entirely different from that of rhombic and monoclinic sulphur to each other or to the liquid. The two crystalline varieties have a transition temperature, about 96° (95.4°, Reicher; 97.6°, Gernez), on each side of which only one of the varieties is capable of permanent existence. The relations of monoclinic to liquid sulphur, and of rhombic to liquid sulphur, are of the same nature, the transition (melting) points being 114.5° (Brodie) for the latter and 120° (Brodie; Gernez gives 117.4°) for the former. The liquids produced by melting these varieties are identical. Amorphous sulphur, on the other hand, is formed when liquid sulphur is heated, and apparently in increasing proportion as the temperature rises. Berthelot<sup>2</sup> states that the maximum proportion is almost reached at 170°, but Küster<sup>3</sup> has shown that the quantity of amorphous sulphur is much greater at 448° than at temperatures below 200°. Even at this temperature, however, the proportions he found did not exceed 34.2 per cent. It seems therefore certain that both bodies coexist throughout a wide range of temperatures. These facts suggest<sup>4</sup> that the bodies are chemically different and exist in equilibrium with one another. The experiments of Küster and Schaum show that the proportion of amorphous sulphur diminishes when the mass is slowly cooled, although a considerable time is required for the reduction to a minimum. In one experiment of Schaum's<sup>5</sup> the greatest reduction was reached after 45 minutes. The change may be considered, in fact, as a reversible action of medium speed, represented by the equation:



Since the action as written is favored by rise in temperature and reverses on cooling, it is endothermal.

In view of the apparent lack of any definite quantitative knowledge relating to the conditions under which amorphous sulphur is formed and of its relation to liquid sulphur, this work was undertaken. The variability in the freezing-point of melted sulphur, which was traced by Gernez<sup>6</sup> to variations in the temperature to which the

<sup>1</sup> Read in abstract before the Royal Society of Edinburgh, June 2, 1902. This is the first part of an investigation of amorphous sulphur which has not yet been completed. This and the subsequent work on the subject will be published in the *Zeitschrift für physikalische Chemie*.

<sup>2</sup> *Jour. f. prakt. Chemie*, Vol. LXXI, p. 360.

<sup>3</sup> *Zeit. f. anorg. Chem.*, Vol. XVIII, p. 369.

<sup>4</sup> SCHAUM, *Liebig's Annalen*, Vol. CCCVIII, p. 18.

<sup>5</sup> *Loc. cit.*, p. 26.

<sup>6</sup> *Comptes rendus*, Vol. LXXXII, p. 1152. Cf. also BRODIE, *Phil. Mag.* [IV], Vol. VII, p. 439; and SCHAUM, *Liebig's Annalen*, Vol. CCCVIII, p. 20.

material had previously been raised and in the time during which it had been held at such temperatures before freezing was begun, appeared to offer the readiest point of attack. It seemed not unlikely that the temperatures of freezing would be found to depend solely on the proportions of amorphous sulphur surviving at the time of freezing, these being themselves affected in some probably complex way by the previous treatment. In other words, it appeared probable that the amorphous sulphur would be found to act as a foreign dissolved body<sup>7</sup> and to depress the freezing-point of the solvent, the soluble sulphur. It was expected that the depression below some fixed point would be found proportionate to the amounts of the dissolved material according to Raoult's law.

If this anticipation proved correct, the way would be opened up for the determination of the molecular weight of amorphous sulphur in solution in liquid soluble sulphur, the only solvent in which it is known to be freely soluble. With the help of the relation

$$\Delta_M = \frac{0.02 T^2}{q}$$

established by van't Hoff between the molecular depression ( $\Delta_M$ ) and the heat of fusion ( $q$ ) and the melting-point of the solvent on the absolute scale ( $T$ ), the molecular depression can be calculated. Thus by comparison with observed depressions produced by known weights the molecular weight can be ascertained. The results given below show that these expectations were fulfilled.

#### THE EXPERIMENTAL METHOD

The sulphur employed was Kahlbaum's pure sulphur and was further purified by recrystallization from carbon disulphide. It was completely soluble in the cold solvent. About 38 grams were used in each experiment.

An ordinary Beckmann thermometer, afterwards compared with a standardized platinum resistance thermometer,<sup>8</sup> was used in determining the freezing-points. The apparatus consisted of a freezing-tube in which the sulphur was placed, an air mantel, and a bath filled with glycerine. The temperature of this bath could be maintained at any desired point by means of a thermostat and stirring apparatus. The temperature of the bath was adjusted at about 4° below the temperature at which the specimen being used was expected to freeze. Another glycerine bath of higher temperature served for the melting of the sulphur, and the amount of amorphous sulphur was varied by altering the temperature of this bath and the time of exposure to its influence. In some experiments heating over the naked flame was used instead. The quantity of amorphous sulphur produced in the melting bath did not require to be known, as it

<sup>7</sup> Küster (*loc. cit.*, p. 370) has shown that amorphous sulphur is not only dissolved to some extent by carbon disulphide, but may be recovered in part when the solution evaporates spontaneously. It is therefore undoubtedly capable of

existing in solution as a body distinct from ordinary sulphur.

<sup>8</sup> We have to thank Mr. Frank B. Jewett, of the Department of Physics, for carrying out this comparison with great exactness.

continually diminished in amount after the tube was transferred to the freezing bath. During the freezing the fluid was stirred in the usual way. An over-cooling of from  $0.2^{\circ}$  to  $0.4^{\circ}$  occurred before the sulphur began to crystallize. No correction was made for the over-cooling, as calculation shows that even the larger over-coolings involved the immediate crystallization of less than 1 per cent. of the solvent, and the other measurements are probably subject to inaccuracies as great as this would produce. As soon as the temperature had been read, the contents of the tube were poured into a cold evaporating dish, where they immediately solidified. On the following day the mass was pulverized and three portions of about 10 grams each were placed in weighed thimbles of filter paper and were successively extracted with cold carbon disulphide in a Soxhlet apparatus. Five fillings of the apparatus were found sufficient for complete solution of the soluble sulphur. The mean of the three results, which seldom differed from one another by more than 0.5 per cent., was taken as a basis for calculation of the final results.

The postponement of the extraction to the following day was required to enable the amorphous sulphur to solidify. When immediate extraction was attempted in specimens containing more than a small proportion of the amorphous body, oily drops of amorphous sulphur found their way through the thimble in appreciable quantities. The application of a correction for the solubility of the amorphous sulphur is described below.

There seemed to be three possible sources of difficulty in this method. In the first place, it was to be feared that the quantity of amorphous sulphur might fall during the operation of taking the freezing-point so rapidly as to vitiate the results. This fear was not realized, however. When freezing had set in, the thread of mercury in the thermometer rose rapidly and came to rest sharply. It remained steady at the maximum point for a considerable time and then descended slowly in consequence of the increasing concentration of the solution brought about by the freezing out of soluble sulphur. If the amorphous sulphur—the solute—had been passing into soluble sulphur at these freezing temperatures with sufficient speed appreciably to diminish its concentration, the thread would have continued to rise slowly, and no definite freezing-point would have been observable. The fact appeared to be that the reversion to ordinary sulphur was going on too slowly to produce any perceptible effect on the freezing temperature during the one or two minutes at the most which the actual observation of the temperature occupied.

The second possibility was that the amount of amorphous sulphur might have appreciably diminished after the freezing had begun and the temperature had been read, and before the whole mass had solidified and cooled. It is to be noticed, however, that the speed of reversion to soluble sulphur might be conspicuous at  $130^{\circ}$  and become inappreciable at temperatures in the neighborhood of  $117^{\circ}$ , since the speed of chemical changes is very markedly affected by small alterations in temperature. To make sure of the absence of any appreciable loss of amorphous sulphur during the



cooling, however, a series of experiments was made in which, after the freezing-point had been read, the contents of the tube were at once poured onto powdered ice. The amounts of amorphous sulphur were not found to be increased by this treatment. The results were for some reason less regular, however, and so the method described above was preferred.

The statements in handbooks and the opinions of many chemists seem to favor the view that the reversion to soluble sulphur is accomplished within a few hours even at the ordinary temperature. The reports of several observers, however, contradict this flatly. Thus Berthelot<sup>9</sup> states that a large proportion of a pure specimen of amorphous sulphur reverts to ordinary sulphur in a year, but that several years are required for complete reversion. Inasmuch, however, as the presence of much soluble sulphur intimately mixed with the amorphous body might be expected to hasten the reversion, some special experiments were made to test the permanence of amorphous sulphur under the circumstances of our experiments. A specimen containing 2.8 per cent. of amorphous sulphur contained, after being heated for ten hours at 70°, 0.57 per cent., and after being heated for the same time at 100°, no amorphous sulphur. A larger quantity was prepared and kept at the ordinary temperature, which on several occasions between March 5 and July 19 reached 30°. The results were as follows :

Date of Analysis	Found in 10 Grams	Mean
February 1.....	.696 .688	.692
February 10.....	.690 .682	.681
March 5.....	.675 .679	.677
July 19.....	.639 .638	.638

These pairs of measurements illustrate incidentally the close agreement between separate analyses of the same sample, which is attainable by the extraction method.

The third difficulty, the solubility of the amorphous sulphur in carbon disulphide, could be disposed of only by proper correction. It seemed likely that, since during the extraction the system had no opportunity remotely to approach a state of equilibrium with respect to amorphous sulphur, the proportion dissolved would vary with the total amount present. Smaller quantities finely diffused through a large mass of soluble sulphur might be expected to suffer a greater loss through solution. A series of mixtures were therefore prepared by mixing intimately various proportions of the two kinds of sulphur. The amorphous sulphur required for the purpose was made by a treatment in which extraction with carbon disulphide, pulverization, and sifting through a fine sieve were repeated several times in order that entire removal of the

<sup>9</sup> *Comptes rendus*, Vol. XLIV, pp. 313, 373.

soluble variety might be insured. The solubility was determined under the same conditions to which the material was exposed in the main experiments. In the following table the first column gives the percentage of amorphous sulphur originally introduced, the second the quantity (mean of three trials) of amorphous sulphur left after extraction of 10 grams of the mixture with five fillings of carbon disulphide, the third the weight of amorphous sulphur which had been dissolved, the fourth the per cent. of the whole amorphous sulphur which the quantity in the third column represented.

Proportion of $S_{am}$ in Mixture	Weight of $S_{am}$ Left by 10 Grams of Mixture	Weight of $S_{am}$ Dissolved	Per Cent. of $S_{am}$ Dissolved
.02	.174	.026	13
.03	.266	.034	11.33
.04	.358	.042	10.5
.06	.546	.054	9
.16	1.504	.096	6
1.00	9.531	.469	4.7

It will be seen that the amount dissolved was relatively much greater in the more highly diluted mixtures, ranging from 4.7 per cent. in pure amorphous sulphur to 13

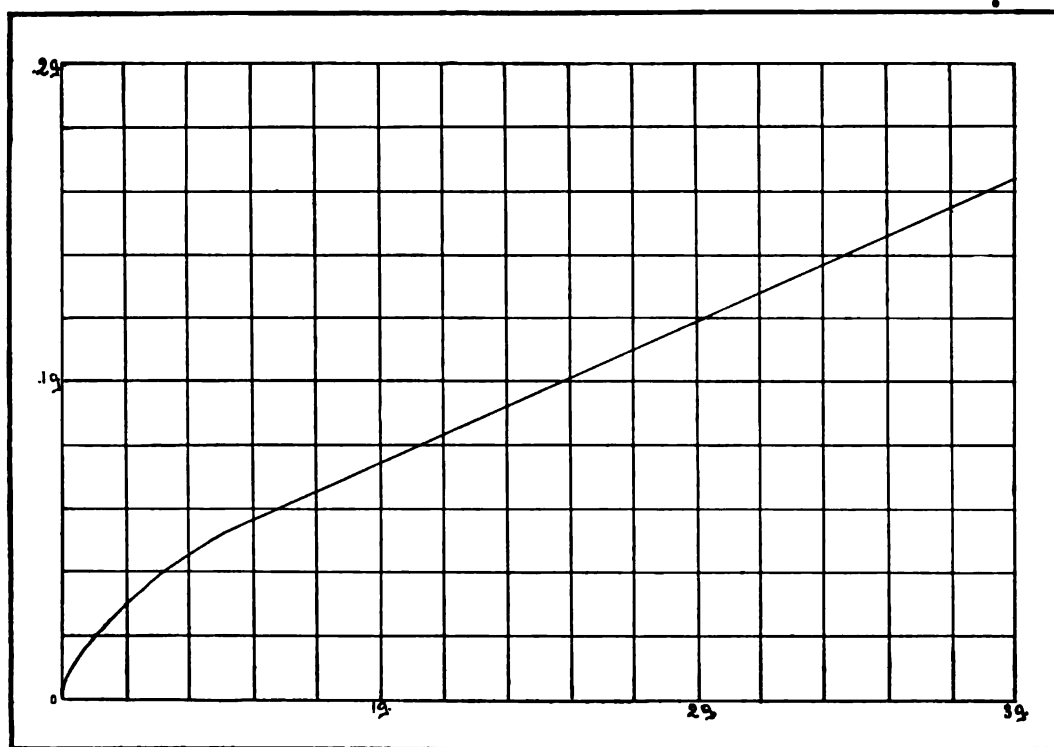


FIG. 1

per cent. in a 2 per cent. mixture. Inasmuch as the most careful incorporation of the ingredients could not give so complete a dissemination of the amorphous sulphur in these artificial mixtures as probably existed in the mixtures produced by heating, it is likely that the solubilities appearing in this table are all somewhat lower—especially in the case of the greater dilutions—than those which natural mixtures of the same concentration would have shown. The amounts added by way of correction to the quantities of amorphous sulphur actually found in the freezing-point experiments were therefore probably somewhat too small.

The curve (Fig. 1) shows the mode in which the correction was applied. The weights of amorphous sulphur left in 10 grams of the mixture (column 2) are laid off along the horizontal axis, and the corresponding quantities of dissolved amorphous sulphur along the vertical axis. Thus, the 10-gram lot being the basis of all extractions, the amount to be added to the mean quantity of amorphous sulphur found in each specimen could be read directly from the curve. The final results are expressed in grams of amorphous sulphur dissolved in and found with 100 grams of soluble sulphur.

#### FREEZING-POINTS AND CORRESPONDING PROPORTIONS OF AMORPHOUS SULPHUR

The accompanying table gives the results, and shows all the measurements of the two series. Experiment No. 9 of the first series must be excluded from subsequent

No. of Experiment		$t$	$S_{am}$
First series—	1.....	118.35	0.711
	2.....	118.29	0.724
	3.....	118.27	0.738
	4.....	118.245	0.761
	5.....	118.22	0.810
	6.....	118.12	0.842
	Mean 1-6....	118.25	0.764
	7.....	117.10	1.546
	8.....	116.20	2.251
	(9).....	(116.00)	(1.705)
	10.....	115.24	3.100
Second series—	11.....	114.43	3.716
	1.....	117.33	1.467
	2.....	116.53	1.932
	3.....	114.84	3.378
	4.....	113.90	4.086
	5.....	112.45	5.291

consideration, as, owing to some error not recognized at the time, the proportion of amorphous sulphur appears to be much less than that found in a specimen of higher melting-point. The second series includes four measurements. After it was completed, a number of efforts were made to secure observations in which larger amounts

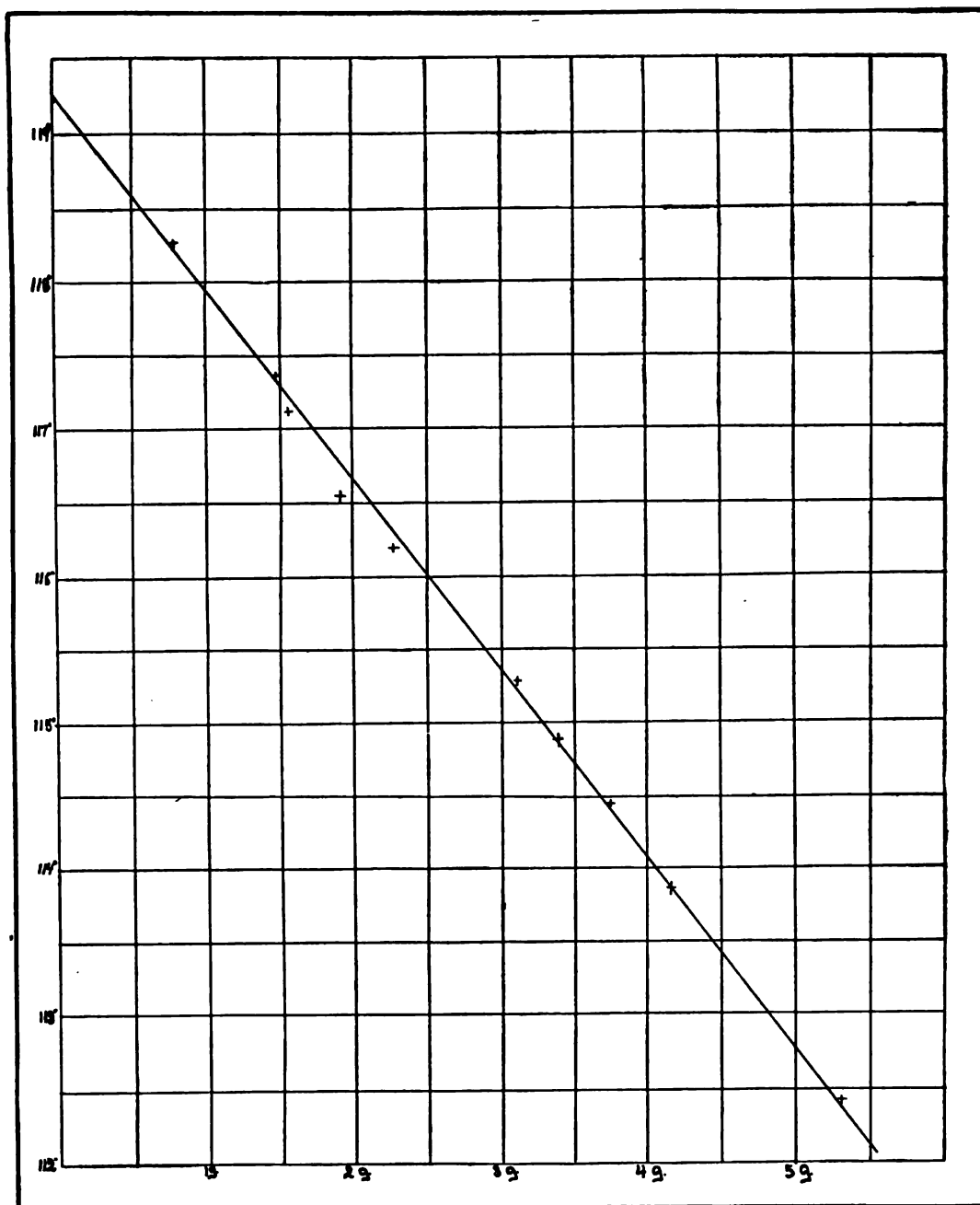


FIG. 2

of amorphous sulphur survived and correspondingly lower freezing-points were attained. Only one of these was successful, and has been added as Experiment 5.

The deliberation with which the cooling had to be conducted to avoid overshooting the freezing-point, and local solidification, was in all cases inevitably accompanied by disappearance of the greater part of the amorphous body. In the second column are given the freezing-points ( $t$ ), and in the third the quantities of amorphous sulphur found ( $S_{am}$ ). The latter are corrected as above and calculated so as to show the proportion dissolved in 100 grams of soluble sulphur.

For the better comprehension of their significance the results of both series are represented graphically in Fig. 2, in which the ordinates represent the freezing-points and the abscissæ the corresponding proportions of amorphous sulphur. The fact that the points all lie upon a straight line shows that the depressions are proportional to the amounts of the dissolved body, as is the case in all true solutions. It will be seen also that the line produced backward cuts the temperature axis at  $119.25^\circ$ , which is the freezing-point when the quantity of the dissolved body is zero. This is, therefore, the melting-point of pure monoclinic sulphur.<sup>10</sup>

#### THE MOLECULAR WEIGHT OF AMORPHOUS SULPHUR

The only previous attempt to determine this constant<sup>11</sup> led to the conclusion that it was the same as for the other varieties of sulphur. But the method was open to many objections, and the conclusion highly uncertain. Phosphorus was used as solvent, and the two crystalline and the amorphous varieties of sulphur were found to cause equal depressions in its freezing-point. Calculation from the heat of fusion gave a value which led to the formula  $S_{10}$ . The solutions violated an essential restriction in work of this kind, being too concentrated (14 and 17 per cent.). The solvent was unfortunately chosen in view of the tendency of the two bodies to interact chemically. And, finally, it was more than doubtful whether the amorphous sulphur really dissolved as such at all. Common sulphur being by far the stabler form at  $27^\circ$  (the temperature of Gloss's experiments), we should feel tolerably certain, unless the contrary was proved, that a solution of common sulphur was formed.<sup>12</sup> Nothing short of recovery of the amorphous sulphur after the experiment—and this was not attempted—could have removed the two last objections.

In the following table will be found the results of the two series of measurements, expressed so as to exhibit, in the first column, the depression below  $119.25^\circ$  ( $\Delta_{obs}$ ), and in the second column the depression which would be produced by 32

<sup>10</sup> Gernæ gives  $117.4^\circ$  and Brodie  $120^\circ$  as the melting-point of monoclinic sulphur. In connection with experiments to be described in the second paper, we have succeeded in preparing melted sulphur almost free from the amorphous variety, and in observing its freezing-point. Two observations gave  $119.165^\circ$  and  $119.17^\circ$ . These samples, when cooled quickly and extracted with carbon disulphide, left only the merest trace of amorphous sulphur. The amount was far too small to permit the application of any correction. According to our curve,  $119.17^\circ$  would correspond to the presence of 0.05 grams of amorphous sulphur

in 100 grams of soluble sulphur. These observations may be taken as confirming strongly the above conclusion, that the true melting-point of the pure body is  $119.25^\circ$ .

<sup>11</sup> GLOSS, *Jour. Phys. Chem.*, Vol. II, p. 421.

<sup>12</sup> The implication that the amorphous sulphur dissolved without change is especially discredited by the results we hope to give in a later paper, in which it will be shown that foreign bodies influence the speed of the action  $S \rightleftharpoons S_{am}$  markedly. (See preliminary note, *Berichte d. deutsch. chem. Gesell.*, Vol. XXXV [1902], Ferienheft.)

grams, one atomic weight, of amorphous sulphur dissolved in 100 grams of soluble sulphur ( $\Delta_{at}$ ). The first value of the first series is the mean of Observations 1 to 6. The remainder are derived from the other observations in the order of the former table, with omission of No. 9:

FIRST SERIES		SECOND SERIES	
$\Delta_{obs} = 119.25 - t$	$\Delta_{at}$	$\Delta_{obs} = 119.25 - t$	$\Delta_{at}$
1.00	41.9	1.92	41.9
2.16	44.7	2.72	45.0
3.05	43.4	4.41	41.8
4.01	41.4	5.35	41.9
4.82	41.5	6.80	41.1
	Mean 42.6		Mean 42.3

The atomic depression should be constant, and it will be seen that the means of the two series agree within the limit of error of the experimental method. The mean value of the atomic depression for all the experiments is  $42.5^\circ$ . The molecular depression is calculated by means of van't Hoff's formula:

$$\Delta_m = \frac{0.02 T^2}{q}$$

where  $T$  is the freezing-point of the solvent on the absolute scale ( $119.25^\circ + 273^\circ$ ), and  $q$  is the heat of fusion of the solvent ( $9.368^{13}$ ). From these data  $\Delta_m = 328$ . Dividing this by the atomic depression, the quotient 7.75 represents the number of atoms in the molecule. Since, as we have already remarked, the quantities of amorphous sulphur are certainly somewhat underestimated, and the atomic depression is, therefore, certainly somewhat too large, we may take it that the molecule contains eight atoms. The equation for the change is thus:

$$S_8 (\text{liquid}) \rightleftharpoons \frac{x}{8} S_8 (\text{amorphous}) .$$

In connection with this result it is interesting to notice what is known about the molecular weight of sulphur under various conditions. Experiments on the vapor density of sulphur made under reduced pressure at temperatures below its boiling-point give values closely approaching  $S_8$ . Thus Bleier and Kohn<sup>14</sup> found a maximum of 7.85 atoms in the molecule; Biltz and Preuner,<sup>15</sup> by the Dumas method, 7.23. In each case the results showed that the maximum toward which the results were tending was not greater than 8. By observations of the freezing-point and boiling-point

<sup>13</sup> PERSON, *Ann. d. Chim. et d. Phys.* [III], Vol. XXI, p. 295.

<sup>14</sup> *Monatshefte d. Chemie*, Vol. XXI, p. 805.

<sup>15</sup> *Zeit. f. physik. Chem.*, Vol. XXXIX, p. 323.

Beckmann,<sup>16</sup> Hertz,<sup>17</sup> Sakurai,<sup>18</sup> Helff,<sup>19</sup> Orndorff and Terasse,<sup>20</sup> Aronstein and Meihuisen,<sup>21</sup> and Tolloczko,<sup>22</sup> all, with the exception of Orndorff and Terasse,<sup>20</sup> found the formula  $S_8$  for the molecule of soluble sulphur. Pekár,<sup>23</sup> by study of the molecular surface energy of solutions in carbon disulphide and sulphur monochloride, reached the same conclusion.

It appears, therefore, that sulphur vapor at low temperatures, soluble sulphur in solution in many solvents, and amorphous sulphur in solution in liquid soluble sulphur, all have the molecular weight represented by the formula  $S_8$ .

Since the amorphous sulphur is insoluble in common solvents, it at least is a distinct chemical substance, and its formation by the equation given above must consist in the main in a molecular rearrangement. On the other hand, the molecular weight of liquid soluble sulphur ( $S_x$ ) is at present unknown. The fact, however, that it shows the well-known strange changes in viscosity when its temperature is raised (and this, as will be shown in a second communication, even when it contains no amorphous sulphur) seems to show that it undergoes molecular changes aside from the reaction we are discussing, and renders it not improbable that the value of  $x$  at some temperatures may be different from 8.

#### CONCLUSIONS

The chief results of this investigation are:

1. A fairly accurate method of analyzing mixtures of amorphous and soluble sulphur has been worked out.
2. The freezing-point of pure liquid sulphur is found to be  $119.25^\circ$ .
3. The freezing-point of pure melted sulphur is shown to depend solely upon the quantity of amorphous sulphur which it contains, and to be depressed below  $119.25^\circ$  proportionately to the amount of the latter.
4. The molecular weight of amorphous sulphur in solution is found to be  $S_8$ .
5. Amorphous sulphur is thus an isomer of soluble sulphur and exists in chemical equilibrium with liquid soluble sulphur whose molecular weight is unknown.

The influence of foreign substances on the rate of formation of amorphous sulphur, the quantities of the latter existing at various temperatures when equilibrium has been reached, and the whole behavior of fluid sulphur are being investigated.

<sup>16</sup> *Zeit. f. physik. Chem.*, Vol. V, p. 76.

<sup>17</sup> *Ibid.*, Vol. VI, p. 358.

<sup>18</sup> *Jour. Chem. Soc.*, Vol. LXI, p. 1001.

<sup>19</sup> *Zeit. f. physik. Chem.*, Vol. XII, p. 196.

<sup>20</sup> *Am. Chem. Jour.*, Vol. XVIII, p. 173. They found the value  $S_8$ .

<sup>21</sup> *Kon. Akad. Wetensch.*, Amsterdam [VI], Vol. III, p. 3 (July, 1898); abs. in *Zeit. f. physik. Chem.*, Vol. XXX, p. 153.

<sup>22</sup> *Bull. Acad. Sci.*, Cracovie, 1901, p. 1; *Centralblatt*, 1901, Vol. I, p. 989.

<sup>23</sup> *Zeit. f. physik. Chem.*, Vol. XXXIX, p. 448.

## **NEW INSTRUMENTS OF PRECISION**





# NEW INSTRUMENTS OF PRECISION FROM THE RYERSON LABORATORY

R. A. MILLIKAN

## I. A SUBSTITUTE FOR THE ATWOOD'S MACHINE

SINCE Galileo's time countless devices have been brought forward as substitutes for his marble and inclined plane in the study of the force of gravity. Many of these have been faultless in principle, but few have been satisfactory in practice. From the standpoint of general utility in an elementary or intermediate laboratory, perhaps the Atwood's machine furnishes the best possible example of a conspicuous failure. The more elaborate forms of this instrument, such as are provided with accurate chronometric appliances, are too expensive and complicated for general laboratory use, while the simpler forms are usually incapable of yielding results of even moderate accuracy; and that for the very obvious reason that the time measurement is a rough ear-estimate.

An instrument which is neither expensive nor complicated, and yet which, in the hands of average students, is capable of furnishing accurate and direct information regarding the most fundamental of the mechanical laws, those of gravity and inertia, is one of the most imperative needs of a physical laboratory. It is thought that these requirements are admirably met by a piece of apparatus which has been gradually worked out at the Ryerson Laboratory during the last four years.

The tuning-fork has always been recognized as one of the most accurate of time measurers, and it is now more than half a century since the idea occurred to both von Babo and Laborde to study the acceleration of gravity by means of the trace which is left upon a smoked glass plate when it falls past a vibrating tuning-fork provided with a stylus. To this day this device in its primitive form is used in some laboratories for the approximate determination of  $g$ . It lacks in accuracy because of the necessary shortness of the falling glass plate and the consequent shortness of the trace.

The obvious advantage of dropping a vibrating fork past a long stationary plate does not appear to have been turned to account until some four years ago. At this time Professor Stratton, then at the University of Chicago, devised a simple machine by means of which the trace of a vibrating fork, carried by a frame which fell without appreciable retardation down suitable guides, was obtained upon a smoked glass plate as much as five feet long. This instrument has undergone considerable modifications at the hands of the writer and also of Mr. Gaertner, the instrument maker, until it is now capable of being used, not only for the study of the motion of a freely falling body, but also for the verification of Newton's second law of motion, *i. e.*,  $f \propto ma$ , or, in fact, for any of the purposes to which an Atwood's machine can be applied.



FIG. 1

The accompanying cut (Fig. 1) shows the final form which the instrument has taken. When it is used for the study of free fall, the cord and balancing weights are removed, the instrument is leveled with a plumb, the frame and fork are carried to the top of the guides, where they are held in place by a special catch to be presently described; the glass plate is blackened and adjusted so that the stylus rests lightly upon the blackened surface; then a lever is thrown which simultaneously releases the frame and sets into vibration the fork. The falling body writes its trace upon the full length of the blackened plate, and is finally brought to rest without appreciable jar in the dash pots at the base of the instrument. The plate is then removed, and the successive distances corresponding to a given number of waves are measured.

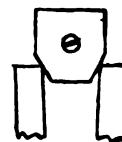


FIG. 2

The method of releasing and exciting the fork is perhaps deserving of especial notice. In the original instrument the prongs were set in motion by slipping from the beveled edges of a stationary steel piece, as shown in Fig. 2. This arrangement was abandoned because, since one prong is always freed slightly in advance of the other, the falling body, instead of descending freely, bounded more or less from side to side as it passed down the ways. The release shown in Fig. 1 and enlarged in Fig. 3 was designed especially to avoid this difficulty. The prongs *p* of the fork are slightly notched at the top, so as to catch over the projecting rim on the lower end of the steel rod *a*. When the lever *b* is thrown to the left, the motion of the collar *c* along the incline *ef* draws up the rod *a*, and thus releases and excites the fork. With this arrangement it will be observed that the prongs are held snugly against the cross-bar *AB* until the instant of release. In order to facilitate getting the fork into position, as shown in the figure, the lower end of *a* is made elliptical. Turning the lever *d* through  $90^\circ$

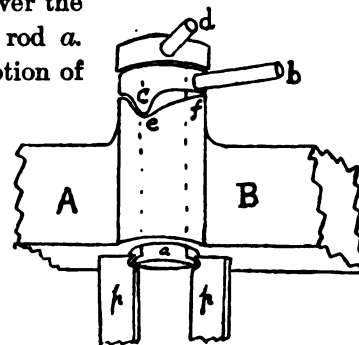


FIG. 3

brings the small axis of the ellipse parallel to the line connecting the prongs. The fork is then pushed up against *AB*, the lever *d* turned back, and the prongs thus spread as the minor axis is replaced by the major.

When it is desired to use the instrument as an Atwood's machine, *i. e.*, to study the case of retarded fall, a cord is attached to the frame, carried over the light aluminum pulley, fastened at the other end to weights of any desired magnitude, and a trace taken precisely as in the case of free fall. The wheel is so light and the frame and fork so heavy (about 1 kg.) that the moment of inertia of the wheel may be left wholly out of account. The adjust-

able platforms shown in Fig. 1 are in reality unnecessary additions, since a study of the trace furnishes the space traversed during any time, the velocity at any instant, or the acceleration at any point. Nevertheless, it might sometimes be desirable to measure *directly* the velocity at some definite point in the fall. This can be done by picking up from the adjustable platform at the desired instant sufficient additional weights to reduce the moving force to zero. The following is an average set of observations upon free fall. It shows about what may usually be expected of the instrument.  $S$  represents the number of centimeters covered by a group of ten waves near the top of the trace,  $S_2$  the space covered by the succeeding group of ten waves; etc., etc.

	Accel.		Accel.
$S_1 = 1.37$	2.18	$S_6 = 10.00$	2.18
$S_2 = 3.55$		$S_7 = 12.18$	
$S_3 = 5.7$		$S_8 = 14.32$	
$S_4 = 7.88$		$S_9 = 16.52$	
$S_5 = 10.00$		Mean	2.165

The uniformity of the acceleration of gravity is nicely shown by the last column. It is chiefly this which the experiment is designed to bring out. Nevertheless, the friction of the guides is so slight that the error in the actual value of  $g$  is scarcely greater than that usually given by direct methods. Thus the rate of the fork was 211 per second. Hence

$$g = 2.165 \times \left(\frac{211}{10}\right)^2 = 963 \text{ cm.}$$

The error in this case, viz., 1.7 per cent., is greater than that sometimes found. It represents a fair average value.

The following is a record of a test made with the pulley and balancing weights:

Weight of frame and fork ( $W$ ) = 1015  
Moving force in gm. ( $m_1$ ) = 450

FIRST TRACE		SECOND TRACE	
Spaces	$a_1$	Spaces	$a_2$
$S_1$ 6.76	1.54	6.10	1.50
$S_2$ 8.30		7.60	
$S_3$ 9.89	1.59	9.15	1.55
$S_4$ 11.38	1.49	10.70	1.55
$S_5$ 13.20	1.82	12.25	1.55
$S_6$ 14.55	1.35	13.82	1.57
$S_7$ 16.20	1.65	15.38	1.56
Means	1.573		1.546
Final mean			$a_1 = 1.560$

$$\frac{f_1}{f_2} = \frac{m_1}{m_2} = .473$$

$$\frac{[W + (m - m_1)] \times a_1}{[W + (m - m_2)] \times a_2} = .468$$

Balancing weight ( $m$ ) = 1013  
Moving force in gm. ( $m_2$ ) = 950

FIRST TRACE		SECOND TRACE	
Spaces	$a_1$	Spaces	$a_2$
$S_1$ 7.40	4.75	4.80	4.95
$S_2$ 12.15		9.75	
$S_3$ 16.95	4.80	14.50	4.75
$S_4$ 21.85	4.90	19.35	4.85
$S_5$ 26.80	4.95	24.30	4.95
	....	29.00	4.70
Means	4.85		4.84
Final mean			$a_2 = 4.845$

Per cent. of error = 1.1.

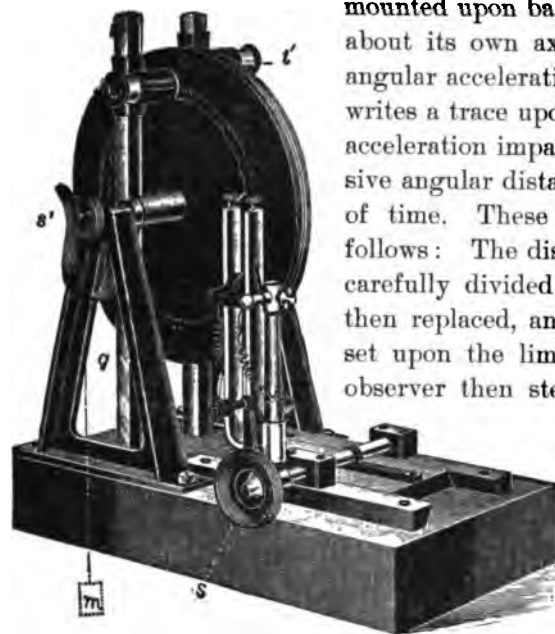
II. A "MOMENT OF INERTIA" MACHINE<sup>1</sup>

FIG. 4

Fig. 4 shows a recently perfected instrument for the study of moment of inertia. The concept of "moment of inertia" is one which the ordinary student has so much difficulty in comprehending that the introduction into a laboratory course of its direct determination from a measurement of the angular acceleration produced by a known moment of force has been found of no small value. A disc weighing several kilograms is mounted upon ball bearings so as to rotate with very little friction about its own axis. While a weight  $m$  imparts to the disc an angular acceleration  $\alpha$ , an electrically driven fork of known period writes a trace upon the blackened face of the disc. The angular acceleration imparted to the disc is found by subtracting successive angular distances traversed during successive equal intervals of time. These angular intervals are obtained in degrees, as follows: The disc is first removed from the frame and the trace carefully divided into groups of, say, fifty waves. The disc is then replaced, and the cross-hairs of the low-power telescope  $t$  set upon the limiting mark of the first group of waves. The observer then steps around to the telescope  $t'$  and takes the reading of the cross-hairs of this telescope upon the circular scale which is graduated upon the unblackened face of the disc. He next rotates the disc until the other limiting mark of the first group of waves comes underneath the cross-hairs of telescope  $t$ , and again takes a reading by means of  $t'$ . The difference between these two readings is the angular

space traversed during the time of fifty vibrations of the fork.

The rack and pinion shown in the figure permit the fork to be moved in a direction parallel to the face of the disc while the trace is being made. This motion prevents the overlapping of the traces which correspond to successive revolutions. The thumb-screw  $s'$  shifts the whole disc in the direction of the axis and thus makes it easy to secure a suitable pressure of the stylus against the blackened face of the disc. The average performance of the instrument is shown by the following examples, in which  $M$  represents the mass of the disc,  $R$  its radius, and  $I$  its moment of inertia; the "calc'd  $I$ " is the quantity  $\frac{MR^2}{2}$ .

<sup>1</sup> The design of the instrument is due to Professor Stratton, Mr. Hobbs, Mr. Gaertner, and the writer.

## FIRST DETERMINATION

 $m = 100$  gm. Fork rate = 150

Angle Reading	Angular Spaces	Angular Acc's
63.25		
102.48	39.23	38.29
180.00	77.52	36.98
294.50	114.50	36.75
445.75	151.25	37.80
634.80	189.05	36.90
860.75	225.95	

Mean acceleration = 37.34

 $R = 10$  cm. $\therefore I = 157080$  $M = 3120$  g. $\therefore$  calc'd  $I = 156000$ 

## SECOND DETERMINATION

 $m = 200$  gm. Fork rate = 150

Angle Reading	Angular Spaces	Angular Acc's
15.53		
65.25	49.72	71.93
186.90	121.65	71.15
379.70	192.80	70.90
643.40	263.70	71.30
978.40	335.00	.....
.....	.....	.....

Mean acceleration = 71.32

 $\therefore I = 154950$ Mean  $I = 156015$ Per cent. of error in observed  $I = .01$ 

## III. A YOUNG'S MODULUS APPARATUS

Most Young's modulus devices which are simple from the point of view of manipulation are woefully inaccurate. The instrument shown in Fig. 5 has recently been designed with a view to combining accuracy with suitability to the needs of general laboratory work. The upper end of the wire whose coefficient of elasticity is sought is firmly clamped in the chuck at  $a$ . At  $b$  it is gripped by a second chuck which is set into a cylindrical brass piece. This cylinder passes, with very little play, through a circular hole in the cross-piece  $n$ ;  $n$  is rigidly clamped to the upright rods  $R$  and  $R'$ , and carries a small horizontal table upon which rest the front feet of an optical lever  $m$ . In order to prevent any easy displacement of the lever, these feet are set in a shallow groove. The rear foot of the lever rests upon the face of the chuck  $b$ . At a distance of twelve or thirteen feet from the instrument is placed a telescope and scale for reading the magnified elongations produced by the successive addition of weights to  $P$ . With kilogram weights and a wire of 0.5 mm. diameter the elongation per kilo-

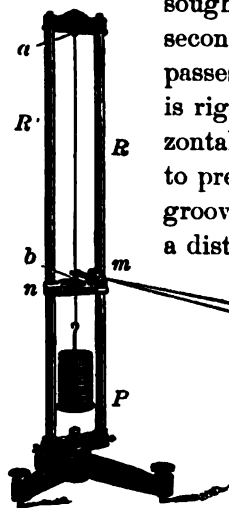


FIG 5

gram observed upon the scale is from 1 to 2 cm., and the divergences from the mean elongation seldom amount to more than 1 per cent. The lever arm of the optical lever can easily be measured to  $\frac{1}{4}$  per cent. The following is an average set of readings:

gram observed upon the scale is from 1 to 2 cm., and the divergences from the mean elongation seldom amount to more than 1 per cent. The lever arm of the optical lever can easily be measured to  $\frac{1}{4}$  per cent. The following is an average set of readings:

Weights in Kilograms	Read- ings	Differ- ences		Weights in Kilograms	Read- ings	Differ- ences	
1	23.32	1.92		7	12.15	1.85	
2	21.40			6	14.00		
3	19.55	1.85		5	15.82	1.88	
4	17.70			4	17.7		
5	15.85	1.85		3	19.55	1.85	
6	14.00			2	21.40		
7	12.15	1.85		1	23.30	1.90	

## IV. A VAPOR-TENSION DEVICE

The laws and phenomena of saturated vapor ordinarily receive in laboratory courses in physics much less attention than they deserve, and this is no doubt partly due to the unsatisfactory nature of the appliances usually used for measuring the tension of saturated vapors. No objection is to be urged in theory against the usual method of the barometer column and water-jacket. It is very unsatisfactory in ordinary practice; first, because of the difficulty of attaching a suitable water-jacket to the top of a barometer column; second, because of the necessity of using a cathetometer for taking the level readings; and, third, because of the impossibility of maintaining a uniform temperature in a long water-jacket at a temperature considerably above that of the room. For practical purposes the method is limited, therefore, to temperatures which differ little from that of the room. Similarly, the observation of the variation of the boiling-point with pressure is limited to a temperature range of a few degrees on either side of  $100^{\circ}$ .

By means of the device shown in Fig. 6, the pressure temperature curve of saturated water vapor can be conveniently, quickly, and accurately followed throughout the whole range from  $0^{\circ}$  to  $110^{\circ}$  C. It is thus made clear that the so-called boiling-point curve and the saturated vapor curve are one and the same.



FIG. 6

The long arm of the apparatus is an ordinary barometer, filled in the usual way through an opening at *c*. Enough mercury is used to fill the bulb *B* half full. A layer of water is then introduced above the mercury and boiled until the air is all driven from the bulb, when the opening at *c* is sealed off. The bulb is then mounted as in the figure and immersed in a simple beaker of water whose temperature can be varied between  $0^{\circ}$  and  $100^{\circ}$ . The vapor tension at any temperature is, of course, the difference between the levels of the mercury in the bulb and in the tube after the latter level has been corrected for the capillary depression of mercury. This difference is read upon an ordinary meter stick. The addition of salt to the water enables the observations to be carried a few degrees above  $100^{\circ}$ .

**ON THE “BECKMANN REARRANGEMENT”**

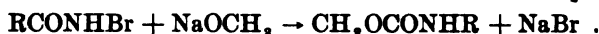
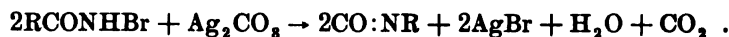




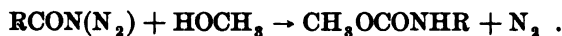
# ON THE "BECKMANN REARRANGEMENT"<sup>1</sup>

JULIUS STIEGLITZ

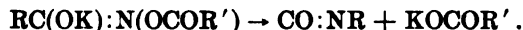
THE molecular rearrangement which has come to be known as the "Beckmann rearrangement" has been produced in three classes of nitrogen derivatives: the acid halogen amides,  $\text{RCON}^*\text{H}(\text{Hal})$ , the oximes of ketones, aldehydes, acid salts, amides, and esters,  $\text{RR}'\text{C}:\text{N}^*\text{OH}$ ,<sup>2</sup> and the acid azides,<sup>3</sup>  $\text{RCON}^*(\text{N}_2)$ . In every case the essential point of the rearrangement is that the alkyl (aliphyl) group, R, is made to leave the carbon atom which holds it in the original substance and is then found joined to the nitrogen atom, N\*, derivatives of an alkyl (aliphyl) amine being formed. The following instances may be used in illustration: acid bromamides on treatment with silver carbonate,<sup>4</sup> alkalies,<sup>5</sup> or sodium methylate<sup>6</sup> produce cyanates, carbamates and urethanes, respectively:



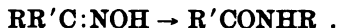
Acid azides, on warming with alcohols, give urethanes:<sup>7</sup>



The dry alkali salt of an acyl hydroxamic acid breaks down spontaneously, yielding an isocyanate:<sup>8</sup>



Ketoximes, on treatment with certain dehydrating reagents, give acid amides:<sup>9</sup>



The determination of the essential cause and mechanism of these rearrangements is a problem of deepest interest which still awaits its final solution, although it has already been the subject of some investigation and speculation. In a paper published six years ago the author developed a theory as to the nature and cause of the rearrangement and outlined certain points of attack which promised to lend themselves to

<sup>1</sup> A first paper was published in the *Am. Chem. Jour.*, Vol. XVIII (1896), p. 751.

<sup>2</sup> R' may be hydrogen, an alkyl, an aliphyl, an alkoxy group, a metaloxyl group, OMe', or an amide group.

<sup>3</sup> An alkyl azide, benzyl azide, also has been found to react in the same fashion.—CURTIUS, *Jour. f. prakt. Chemie*, Vol. LXIII, p. 429.

<sup>4</sup> HOFMANN, *Berichte d. deutsch. chem. Gesell.*, Vol. XV, p. 412.

<sup>5</sup> HOOGWERFF AND VAN DORP, *Recueil d. trav. chim. d. Pays-Bas*, Vol. VIII, p. 200.

<sup>6</sup> LENGFELD AND STIEGLITZ, *Am. Chem. Jour.*, Vol. XV, pp. 215, 504.

<sup>7</sup> CURTIUS, *Berichte d. deutsch. chem. Gesell.*, Vol. XXVII, p. 778; *Jour. f. prakt. Chemie*, Vol. L, p. 289.

<sup>8</sup> Unpublished results of Dr. L. W. JONES, of this University. Vide also HANTZSCH, *Berichte d. deutsch. chem. Gesell.*, Vol. XXVII, p. 1256, and THIELE, *Liebig's Annalen*, Vol. CCCIX, p. 189.

<sup>9</sup> BECKMANN, *Berichte d. deutsch. chem. Gesell.*, Vol. XIX, p. 988.

the experimental investigation of the correctness of the theory. The work has progressed very slowly, largely because an important part was undertaken in collaboration with Mr. E. E. Slosson, of the University of Wyoming, and the latter was able to devote only his summer vacations to the prosecution of the investigation. The work with Mr. Slosson having now reached a certain stage of completeness, it is proposed in this paper to consider the present aspect of our knowledge of the "Beckmann rearrangement," on the basis of all the general facts known in connection with it and particularly on the basis of the results obtained with Mr. Slosson, as well as in shorter investigations carried on with Drs. O. Folin, H. H. Higbee, H. N. McCoy, J. H. Ransom, and Elizabeth Jeffreys.

It may be said that any interpretation of the "Beckmann rearrangement" must take cognizance of these fundamental points:

1. The nature of the rearrangement of the *three* classes of nitrogen derivatives described above seems to be substantially the same. This follows from the strictly analogous character of the initial substances, from the adaptation of the reagents used in effecting the rearrangement to the variation in the elements combined with the nitrogen in the compounds used, and from the frequent identity and otherwise close relationship of the products obtained from all three classes of compounds.

2. The rearrangement is produced by reagents which stand in the most intimate relation to the elements combined with the nitrogen in each of the classes of nitrogen compounds under investigation. In the case of the acid halogen amides,  $\text{RCONH}(\text{Hal})$ , the effect is produced by substances which have the greatest affinity for the halogen acids—for instance, the alkalies and silver carbonate—and salts of these acids are always obtained in the reaction. The halogen atoms go from the condition of "hypo-acid" radicals<sup>10</sup> into simple negative ions and their salts. Van Dam<sup>11</sup> has shown, in fact, that the velocity of the rearrangement is proportional to the concentration of hydroxyl ions in the solution. The rearrangement of the oximes of ketones, aldehydes, and acid esters is effected by dehydrating reagents, water, or rather derivatives of water, resulting;<sup>12</sup> the rearrangement of the acylhydroxamic acids,  $\text{RCONH}(\text{OCOR}')$ , on the other hand, is again caused by alkalies, salts of the organic acids,  $\text{HOCOR}'$ , being formed. Finally the rearrangement of the acid azides,  $\text{RCON}(\text{N}_3)$ , requires no such reagent; it is produced essentially by the action of heat, nitrogen being evolved in all such cases.

3. When in the case of the acid bromamides,  $\text{RCONH}^*\text{Br}^*$ , the possibility of the loss of hydrobromic acid by the union of the atoms marked \* in the formula is cut off by replacing the hydrogen atom by an alkyl or an aliphyl group, the tendency to undergo the "Beckmann rearrangement" has been found also to be entirely suppressed. Mr. Slosson and I have shown this to be the case for the acyl alkyl halogen

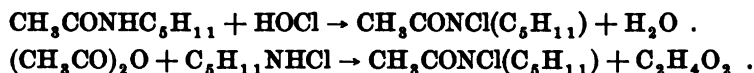
<sup>10</sup> SELIWANOW, *ibid.*, Vol. XXV, p. 3620.

<sup>11</sup> *Recueil d. trav. chim. d. Pays-Bas*, Vol. XVIII, p. 408; Vol. XIX, p. 318. In the same investigation Van Dam

proved that the reaction is a monomolecular one, the rearrangement, therefore, *intramolecular*.

<sup>12</sup> *Vide Am. Chem. Jour.*, Vol. XVIII, p. 754.

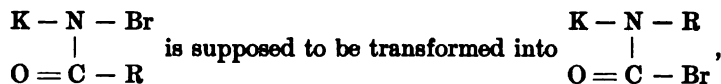
amines,<sup>13</sup> whose constitution we determined by the preparation of identical substances by the action of hypochlorous acid on acetylamine, on the one hand, and of acetic acid anhydride on chloramine, on the other hand:



Chlorimidoesters,  $\text{RC}(\text{:NCl})\text{OR}'$ , also have lost the tendency<sup>14</sup> to suffer the molecular rearrangement so characteristic of the parent substances, the acid halogenamides. In like fashion, according to an investigation carried on by Dr. Higbee, at my suggestion, derivatives of the oximes, in which the possibility of the breaking off of water from a trivalent nitrogen atom is excluded, refuse to undergo the rearrangement when treated with the same dehydrating reagents which are so effective in the case of ketoximes. Benzoyl- $\beta$ -phenylhydroxylamine,  $\text{C}_6\text{H}_5\text{CON}(\text{C}_6\text{H}_5)\text{OH}$ , under such treatment does not give the least discernible trace of diphenylamine, and yet the color tests for this substance belong to the most sensitive known in organic chemistry.

On the basis of these general facts, and with the aid of a number of other experimental facts which will be referred to in the course of the discussion, the most important of the theories advanced in explanation of the "Beckmann rearrangement" may be critically considered.

A. W. Hofmann, whose work on the acid bromamides gave us the first instances of this molecular rearrangement, never suggested any explanation of the important reaction he discovered. Hoogewerff and Van Dorp, who continued Hofmann's work with great skill, advanced the first hypothesis<sup>15</sup> as to the nature and cause of the rearrangement of the acid bromamides, to whose alkali salts they assigned the constitution,  $\text{RCO}(\text{NMe}'\text{Br})$ ; by intramolecular migration the salt



which in turn is decomposed into potassium bromide and an isocyanate. The transposition is caused, in particular, by the tendency of nitrogen to combine with radicals more positive than the halogens. In support of their view they point out that the intramolecular migration must depend largely on the nature of the group, R; when this group becomes more negative, the transposition must become more difficult. Observations made at the time the theory was advanced confirmed this conclusion; the salts of the nitrobenzbromamides were found to be more stable than those of benzbromamide. It was found later, however, by McCoy and myself,<sup>16</sup> that dibromsalicylicbromamide, in which R is the exceedingly negative radical  $\text{C}_6\text{H}_4\text{Br}_2(\text{OH})'$ ,

<sup>13</sup> *Vide* preliminary reports, *Berichte d. deutsch. chem. Gesell.*, Vol. XXVIII, p. 3266; and Vol. XXXIV, p. 1613, and an article which will be published soon in the *Am. Chem. Jour.*

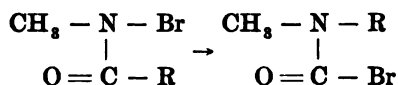
<sup>14</sup> *Ibid.*, and STIEGLITZ, *Am. Chem. Jour.*, Vol. XVIII, p. 751.

<sup>15</sup> *Recueil d. trav. chim. d. Pays-Bas*, Vol. VI, p. 373;

Vol. VIII, p. 173, etc. They were the first to call attention to the similarity of Hofmann's reaction and the "Beckmann rearrangement" of the oximes. *Vide also* GRAEBE AND ROSTOVZEFF, *Berichte d. deutsch. chem. Gesell.*, Vol. XXXV, p. 2747; and HANTZSCH, *ibid.*, Vol. XXXV, p. 3579.

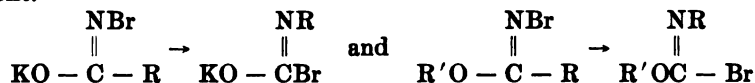
<sup>16</sup> *Am. Chem. Jour.*, Vol. XXI, p. 116.

suffers the rearrangement even below  $5^{\circ}$  in the course of a few seconds, so rapidly that the bromamide could not be isolated. Van Dam<sup>17</sup> in Hoogewerff and Van Dorp's laboratory obtained exactly the same results working at  $-12^{\circ}$ , independently of, and almost simultaneously with, the work done in this laboratory. As to the main point in Hoogewerff and Van Dorp's hypothesis, the theory fails to adapt itself to the entirely analogous behavior of the acid azides and acid bromamides, and to explain the fact that acyl alkyl halogen amines do *not* rearrange.<sup>18</sup> One would expect the intramolecular transposition of the bromine and the alkyl, R, as in



to occur quite as easily as in the case of the alkali salts, (see above), if it is merely a matter of the relative affinity of nitrogen for alkyls and halogens respectively.

The alkali salts of the acid bromamides are now recognized as having a different constitution from that assigned originally to them by Hoogewerff and Van Dorp; the constitution<sup>19</sup> is now known to be  $\text{RC}(\text{OMe}); \text{NBr}$ . But even the esters corresponding to such salts, the chlor- and bromimido esters, do not undergo the "Beckmann rearrangement."



would be strictly analogous, according to the Dutch chemists' interpretation, but the second reaction does not take place. In conclusion, the important function of the reagents which produce the rearrangement of the acid bromamides (see p. 4, 2) seems to find no place in this theory.<sup>20</sup>

<sup>17</sup> *Loc. cit.*

<sup>18</sup> STIEGLITZ AND SLOSSON, *loc. cit.*

<sup>19</sup> See below, p. 13.

<sup>20</sup> In a short note (*Berichte d. deutsch. chem. Gesell.*, Vol. XXV, p. 3579), published after this paper was written, HANTZSCH, while accepting Hoogewerff and Van Dorp's interpretation of the rearrangement, suggests that the rearrangement of the bromamides is caused by an alkali essentially because the latter is necessary to produce a double bond between the acyl carbon and the nitrogen. The constitution of the salts,  $\text{RC}(\text{OMe}); \text{NBr}$ , is said to bring them into closest analogy to the oximes,  $\text{RR}'\text{C}:\text{NOH}$ . After the establishment of the double bond the rearrangement is supposed to proceed by a direct exchange in position between the radical, R, and the bromine atom, Br, of the bromamide salts. This possibility that a double bond might be an essential condition for the rearrangement, and might explain best the rearrangement of the stereoisomeric oximes, was the chief reason for the study of the chloro- and the bromimidoesters,  $\text{RC}(\text{OR}'); \text{NHal}$ , by the author, (*vide* the first paper, *loc. cit.*, and the article by MR. SLOSSON, in the February number (1903) of the *Am. Chem. Jour.*). They were found to be remarkably stable nitrogen halides, which not only do not suffer a "Beckmann rearrangement" spontaneously at ordinary temperatures, as do the salts of

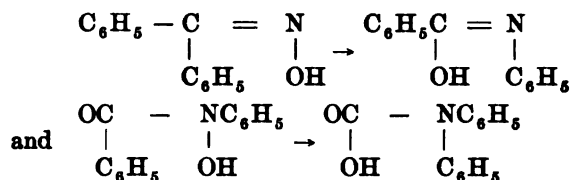
the acid bromamides, but do not rearrange even when heated to  $100-140^{\circ}$ ; in fact, they have never as yet been observed to rearrange under any conditions whatsoever. Stereoisomers do not seem to exist, and preliminary experiments, which will be continued by Mr. L. A. Test, indicate that the acid chloramides produce, on the one hand, with diazomethane the stable chlorimidoesters and with alkalis, on the other hand, the easily rearranged salts.

These facts show that the double bond is, after all, not the vital factor that Hantzsch assumes. The effect of the alkali must be something essentially more than to produce a double bond. The ready and sometimes spontaneous rearrangement of the salts of acyl hydroxamic acids,  $\text{RC}(\text{ONa}); \text{NOAcyl}$ , and the stability of the corresponding esters  $\text{RC}(\text{OR}'); \text{NOAcyl}$ , emphasize the same point; [the acetyl ester of ethyl-syn-benzhydroxamic acid can be heated to  $140^{\circ}$  without change, according to WERNER, *Berichte d. deutsch. chem. Gesell.*, Vol. XXV, p. 41]. Hantzsch's assumption of the intermediate formation of a chlorimide,  $\text{R}_2\text{C}:\text{NCl}$ , in the rearrangement of the ketoximes under the influence of phosphorus pentachloride also appears unwarranted in view of the above facts, and in view of the further fact that other acid dehydrating reagents have the same effect. Finally, acyl azides have no such double bond between carbon and nitrogen, and yet they are rearranged with remarkable smoothness.

Beckmann, the discoverer of the molecular rearrangement of the oximes of ketones and aldehydes, whose name is now used to designate this typical transformation, advanced an explanation<sup>21</sup> of the reaction in 1894, which is very much like the one just discussed. He suggested, rather vaguely, that the rearrangement might consist simply in the transposition of electrically charged radicals or ions, and that the dehydrating reagents employed act chiefly, although not wholly, as catalytic agents giving the necessary impetus for a change from a labile to a stable condition. In the reaction

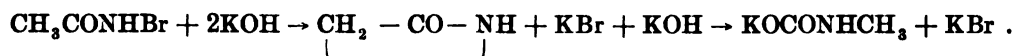


a direct exchange of the phenyl and hydroxyl groups is assumed. The theory is essentially a restatement of obvious results obtained with only one class of compounds, the oximes. Like the previous theory, it fails to explain, on the one hand, the analogous behavior of the oximes, and acid azides, and, on the other hand, the complete difference in behavior of compounds most closely allied to the oximes, such as  $\beta$ -benzoyl- $\beta$ -phenyl-hydroxylamine. There is no evident reason, according to Beckmann's view, why both the following reactions should not occur with equal ease:



and yet the second reaction does not take place even to a slight extent. Finally the characteristic adaptation of the reagent effecting the rearrangement to the class of nitrogen compounds used (see p. 4), is also left practically out of consideration in this interpretation.

In an unpublished investigation carried out at this university in 1896, Dr. Bernhard C. Hesse examined experimentally the reliability of an entirely new explanation of the "Beckmann rearrangement," based on the following views: it was thought possible that the transformation of acetbromamide, for instance, was accomplished through the following stages:



The intermediate formation of a triatomic ring was the basis of the explanation, in analogy to similar known changes of position of alkyls through the intermediate formation of trimethylene rings. Dr. Hesse has kindly consented to let me publish the results of this investigation.

The reliability of the theory was tested by studying the rearrangement product of trimethylacetbromamide,  $(\text{CH}_3)_3\text{C}.\text{CONHBr}$ . An intermediate triatomic ring formation being impossible in this case, at best a four atomic ring could result. In the

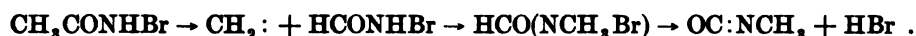
<sup>21</sup> *Berichte d. deutsch. chem. Gesell.*, Vol. XXVII, p. 300.

latter event, isobutylamine,  $(\text{CH}_3)_2\text{CHCH}_2\text{NH}_2$ , would be the final product, while, if no ring formation at all occurred, tertiary butylamine,  $(\text{CH}_3)_3\text{C.NH}_2$ , would be formed. Dr. Hesse identified the salts obtained as those of tertiary butylamine—thus disposing of the possibility that intermediate unstable ring formations of the nature indicated are the cause of the rearrangement.

It may be added that a triatomic ring, as used in the above equation, would, doubtlessly, open between the carbonyl and the imide groups, as all acid amides are saponified in this way.

Only a very brief reference need be made to the views of Freundler.<sup>22</sup> He suggests that the rearrangement of the acid bromamides could be due to a series of reactions, the essential feature of which is the intermediate formation of carbylamine and hypobromite. These views were suggested by the erroneous idea that an isocyanate had never been isolated as an intermediate product in Hofmann's reaction. It is hardly necessary to add that none of Freundler's supposed intermediate products have ever been isolated in the reaction under discussion, and that it would be impossible to bring his view into accord with the behavior of the acid azides and the acyl-hydroxamic acid salts.

The most recent suggestion as to the nature of the "Beckmann rearrangement" has been made, incidentally, by Nef<sup>23</sup> in connection with an exhaustive study of dissociation phenomena of alcohols, ethers, and organic acid salts. Having shown that an alkylidene residue,  $\text{RCH:}$ , or a phenylene group,  $\text{C}_6\text{H}_4:$ , is often broken off from one of the above compounds, and sometimes absorbed again in a different position from that occupied in the original molecule, Nef suggests that the rearrangement of the acid bromamides and of the oximes is due to similar changes. For instance, acetbromamide is supposed to undergo the following series of changes:

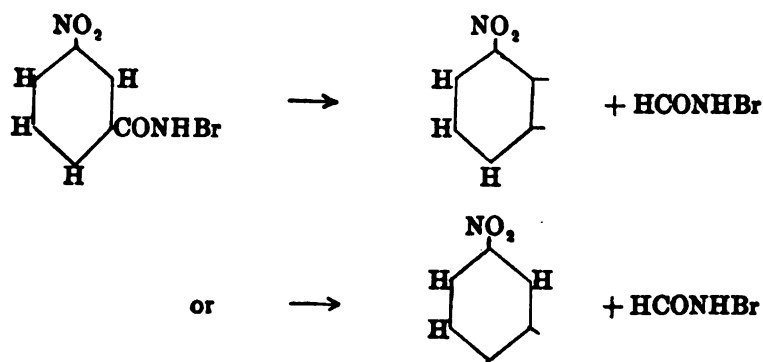


Further close investigation may, possibly, be needed absolutely to prove or disprove the correctness of this interesting view, and with the consent of Dr. Nef experiments lending themselves to a further test of the theory will be carried out. At present, however, it may be said that it cannot be applied directly to the acid azides, and that it leaves quite out of consideration the fact that the nature of the reagents used to effect the molecular rearrangement in each of the three classes of compounds under discussion stands obviously in closest relation to the elements or groups combined with the nitrogen, and not to the alkyl or aliphyl group attached to the acyl carbon atoms (see p. 4). The most important difficulty, in the way of Nef's view is the following: in every product of the "Beckmann rearrangement" known to the author, the nitrogen atom is found to occupy exactly the same position in the alkyl chain or aliphyl nucleus which was held originally by the carbon atom that separated the migrating alkyl or aliphyl group from the nitrogen atom. Thus trimethylacetbromamide,

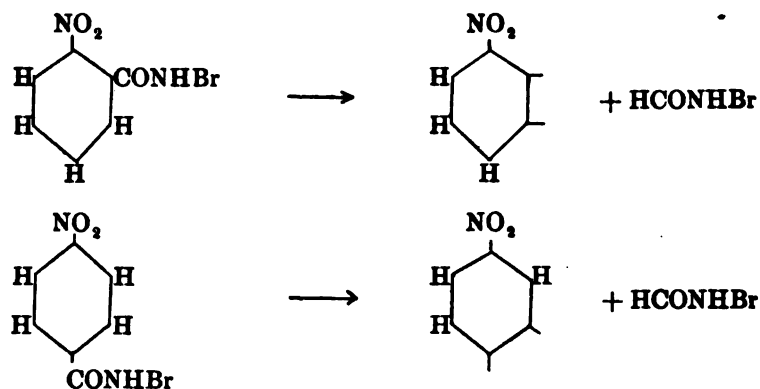
<sup>22</sup> *Bull. soc. chim.*, Vol. XVII, p. 421.

<sup>23</sup> *Liebig's Annalen*, Vol. CCXCVIII, p. 202; Vol. CCCXVIII, pp. 30, 227.

$(\text{CH}_3)_3\text{C}.\text{CONHBr}$ , according to the work of Dr. Hesse reported above, gives tertiary butylamine,  $(\text{CH}_3)_3\text{C}.\text{NH}_2$ , and not isobutylamine,  $(\text{CH}_3)_2\text{CHCH}_2\text{NH}_2$ ; and the three nitrobenzbromamides, para, meta, and ortho, give the corresponding para, meta, and ortho nitranilines or their derivatives.<sup>24</sup> This is not necessary and hardly possible always, according to Nef's view, for derivatives of tertiary and aromatic acids, ketones and aldehydes are supposed to dissociate into olefines or phenylenes,<sup>25</sup> as the methylene dissociation is excluded, and these residues are said to act as the methylene does in the above reaction. Under such circumstances one would have to expect that meta nitrobenzbromamide would give, to a considerable extent, the same aniline derivatives as either ortho or para nitrobenzbromamide; from the reactions



the same intermediate nitrophenylene must result as from either *one or the other* of the following dissociation reactions:



But, as a matter of fact, each nitrobenzbromamide gives exclusively an excellent yield of the corresponding nitraniline or its derivatives.<sup>26</sup> Whether traces or very small quantities of the isomers are also formed remains a question for further

<sup>24</sup> HOOGWERFF AND VAN DORP, *Recueil d. trav. chim. d. Pays-Bas*, Vol. VIII, pp. 193, 198, 199; LENGFELD AND STIEGLITZ, *Am. Chem. Jour.*, Vol. XIV, p. 311; FOLIN, *ibid.*, Vol. XIX, p. 325, etc.; SWARTZ, *ibid.*, Vol. XIX, p. 303;

JEFFREYS, *ibid.*, Vol. XXII, p. 41; STRUVE AND RADHAUSEN, *Jour. f. prakt. Chemie*, Vol. LII, p. 230.

<sup>25</sup> NEF, *loc. cit.*

<sup>26</sup> E. g., *vide* FOLIN, *loc. cit.*

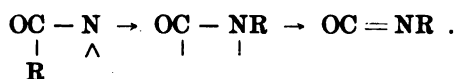


investigation. One would expect, however, not traces, but considerable quantities of identical substances, in view of the above discussion.

The actual dissociation tendency of compounds of the general type,  $RR'NX$ , in which  $X$  is a halogen, a hydroxyl or an acid group,  $RCO$ , as will be shown below, has been observed in a number of instances to be rather toward the breaking off of  $R'X$  from the molecule, leaving most likely an unsaturated imide group with univalent nitrogen,  $RN:$ , as a primary but unisolated product of dissociation. Such a result would be analogous to the alkylidene formations exhaustively studied by Nef. It seems at present, therefore, that the conceptions of "methylene and phenylene dissociations" will aid in clearing up the problem of the "Beckmann rearrangement" rather by analogy than by direct application.<sup>27</sup>

The last interpretation of this molecular rearrangement to be discussed is that proposed by the writer in 1896, the experimental study of which forms the basis of this paper. It may be briefly stated as follows: Starting out from the pronounced similarity in the behavior of the parallel classes of derivatives,  $RCONHBr$ , (I),  $RCNH(OCOR)$ , (II), or  $R_2C(NOH)$ , (II'), and  $RCN(N_2)$ , (III), an initial change which could be common to all three classes of compounds was sought and found in the possibility of all three series of compounds easily yielding an identical intermediate derivative (for instance  $(RCO)N:$ ), with univalent nitrogen; the first class would do so by the initial loss of hydrobromic acid; the second (II), by the loss of an organic acid,  $HOCOR$ , or (II') by the loss of water from salts such as the hydrochlorides,  $R_2CCl.NHOH$ , the result being  $R_2CCl.N:$ ; <sup>28</sup> the third, simply by that loss of nitrogen which is so common with diazo derivatives.

The nature of the reagents used to effect the rearrangement is in complete agreement with this conception, alkalies being used both for (I) and (II), acid dehydrating reagents for (II'), and essentially heat only for (III). It is the intermediate product, such as  $RCO.N:$ , <sup>29</sup> which is considered to suffer the molecular rearrangement proper, owing to the potency of the free valences of univalent nitrogen: for instance,



an isocyanate being formed in this case as the first tangible product.<sup>30</sup>

According to this interpretation, the rearrangement should not be possible if, for instance, in the acid bromamides the loss of hydrobromic acid is prevented by replacing the hydrogen atom by an alkyl or an aliphyl group. The work with Mr. Slosson has proved that such is actually the case, neither series of derivatives of

<sup>27</sup> Vide the note by NEF, *Am. Chem. Jour.*, Vol. XVIII, p. 753.

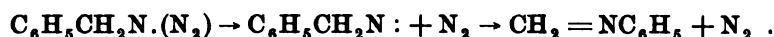
<sup>28</sup> Vide the first paper, *loc. cit.*

<sup>29</sup> TIEMANN, I find, expressed, incidentally, a similar view in regard to one of the above classes of compounds, *Berichte d. deutsch. chem. Gesell.*, Vol. XXIV, p. 4163.

<sup>30</sup> HOFMANN, *Berichte d. deutsch. chem. Gesell.*, Vol. XV, pp. 411, 412; HANTZSCH, *ibid.*, Vol. XXVII, p. 1236; THIELE, *Liebig's Annalen*, Vol. CCCIX, p. 189; DR. L. W. JONES, of the Kent Chemical Laboratory of this University, in a private communication to the author of this paper; vide GRAEBE, *loc. cit.*

the acid bromamides,  $\text{RCO}(\text{NR}'\text{Br})$ , and  $\text{RC}(\text{OR}')\text{NBr}$ , undergoes the change in question. Likewise, it will be remembered, Dr. Higbee found that oximes,  $\text{RCO} \cdot (\text{NC}_6\text{H}_5)\text{OH}$ , which cannot lose water in the way described, show no tendency whatever to undergo the "Beckmann rearrangement." The interpretation is, therefore, in agreement with the three most striking, fundamental characteristics of the rearrangement.

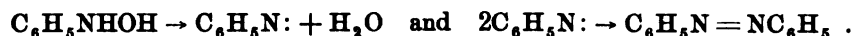
It is also in complete accord with the most recently discovered case of a "Beckmann rearrangement," viz., that of benzyl azide according to:<sup>21</sup>



Not one of the other theories proposed can account for this additional instance of a "Beckmann rearrangement."

In further support of the views expressed, the following instances will be mentioned in which the formation of univalent nitrogen derivatives from the classes of compounds under discussion is clearly indicated:

1. An exhaustive study of the many remarkable reactions of  $\beta$ -phenylhydroxylamine and its homologues has led Bamberger to the conclusion that a phenylimide is the first unisolated product from which many varied derivatives are obtained. For instance, on heating  $\beta$ -phenylhydroxylamine, azobenzene and water are the chief products of decomposition, a reaction which may be interpreted:

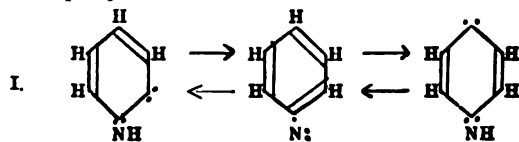


At the same time, the characteristic odor of a carbylamine,  $\text{RNC} :$ , is developed exceedingly strongly. As an ordinary carbylamine can hardly be produced under these circumstances, the suspicion is aroused that the odor is due to small quantities of the analogous phenylimide,  $\text{C}_6\text{H}_5\text{N} :$ , escaping instantaneous condensation.<sup>22</sup>

<sup>21</sup> CURTIUS, *Jour. f. prakt. Chemie*, Vol. LXIII, p. 423; *Berichte d. deutsch. chem. Gesell.*, Vol. XXXV, p. 3229.

<sup>22</sup> *Ibid.*, Vol. XXVII, p. 1550; Vol. XXXIII, p. 3600; Vol. XXXIV, p. 61.

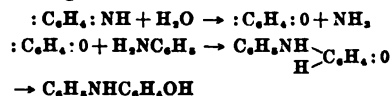
<sup>23</sup> It seems very probable, however, that phenylimide, as soon as formed, goes over, partly, into the isomeric imido phenylenes:



and it remains a question, therefore, to which of these substances the carbylamine odor is really due.

This conception of the partial isomerisation of phenylimide seems to me to give a better interpretation than Bamberger's of all of the reactions of phenyl hydroxylamine discussed and investigated so brilliantly by Bamberger (*loc. cit.*), in which the para (or ortho) position of the phenylimide enters into action. It accounts for the great and varied reactivity of the para position in taking

up, additively, many compounds, such as water, alcohols, anilines, phenoles, etc. It also explains, readily, one reaction, which even Bamberger found particularly peculiar and difficult to understand (*Berichte d. deutsch. chem. Gesell.*, Vol. XXXIV, p. 62), namely, the change of phenyl hydroxylamine into para oxydiphenylamine, on being heated with dilute sulphuric acid; the partial saponification to a phenole derivative, which Bamberger hardly could account for, is accomplished at the imido phenylene stage of the reaction:



(For the formation of aniline see *Berichte d. deutsch. chem. Gesell.*, Vol. XXXIII, pp. 3605-17).

It may be added that with the help of Bamberger's fine work and ideas the conception developed here gives us, at last, also a complete explanation of the "benzidine rearrangements" of the hydrazines. It is noteworthy that these rearrangements are very easily effected by the agency of even dilute acids—a fact readily understood from the following:

2. Chattaway and Orton<sup>36</sup> have found, likewise, that phenylchloramine,  $C_6H_5NHCl$ , breaks down readily into hydrochloric acid and azobenzene, and phenylbromamine still more readily into hydrobromic acid and azobenzene.

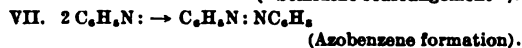
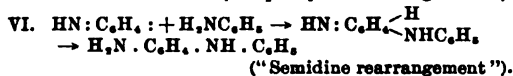
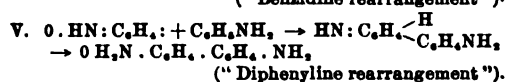
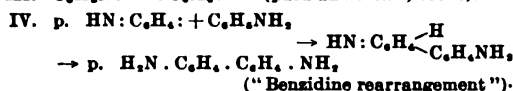
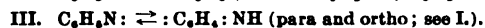
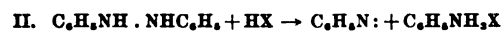
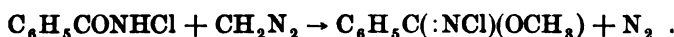
3. The same chemists<sup>36</sup> and Slosson<sup>38</sup> have found that benzoyl phenylchloramine,  $C_6H_5CO.N(C_6H_5)Cl$ , gives, on being heated, some benzoyl chloride<sup>37</sup> (besides forming chiefly p-chlorbenzanilide). In this connection it is interesting to note that the action of aqueous alkali on the acyl aliphyl halogen amines,  $RCO.N(C_6H_5)Cl$ , while producing an acyl anilide as the chief product, was invariably observed to be accompanied by the formation of an unknown compound with the pungent carbylamine odor.<sup>38</sup>

4. By the action of phosphorus pentachloride on benzoyl- $\beta$ -phenylhydroxylamine,  $C_6H_5CO.N(C_6H_5)OH$ , Dr. Higbee has found that some benzoyl chloride is formed and again the inevitable carbylamine odor was most pronounced. Treatment of the same substance with acetic and hydride likewise produced this characteristic odor.

5. The reduction products of the acyl azides also indicate an intermediate formation of the elusive univalent nitrogen derivatives.<sup>39</sup>

6. Curtius<sup>40</sup> deduces the formation of the main decomposition products of benzyl azide from the rearrangement of benzylimide,  $C_6H_5CH_2N$ :

In the first paper on the "Beckmann rearrangement" mention was made of the fact that no determination had been made of the constitution of the metal salts of the acid halogen amides which are formed before the rearrangement is effected.<sup>41</sup> The two forms,  $RCO.N(MeHal)$  (I), and  $RC(OMe)(NHal)$  (II), were considered possible. This determination has since been made by Dr. Ransom and myself<sup>42</sup> and by Hantzsch.<sup>43</sup> Ransom and I have found that diazomethane converts benzchloramide into methyl chlorimido benzoate<sup>44</sup> according to:



We find thus all the interesting products of this important action fully accounted for. These simple interpretations give a further instance of the usefulness of the conceptions of univalent nitrogen derivatives as intermediate products in molecular rearrangements whose course has heretofore always baffled chemists' understanding.

<sup>36</sup> *Jour. Chem. Soc.*, Vol. LXXIX, p. 462. Similar cases are described in the same paper.

<sup>38</sup> *Ibid.*, Vol. LXXV, p. 1053.

<sup>39</sup> *Loc. cit.*

<sup>37</sup> The nature of the corresponding nitrogenous dissociation product has not been determined as yet. The reaction will be studied further.

<sup>38</sup> SLOSSON, *loc. cit.*

<sup>39</sup> STIEGLITZ, *Am. Chem. Jour.*, Vol. XVIII, p. 753.

<sup>40</sup> *Loc. cit.*

<sup>41</sup> STIEGLITZ, *loc. cit.*

<sup>42</sup> Preliminary note, *Berichte d. deutsch. chem. Gesell.*, Vol. XXXIV, p. 1615.

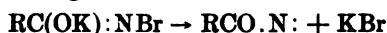
<sup>43</sup> *Ibid.*, Vol. XXXV, p. 228.

<sup>44</sup> The ester was isolated by distillation and characterized by its action on hydroiodic acid and on ammonia. By reduction with hydrogen chloride in ligroin solution it gave the hydrochloride of methyl imido benzoate,  $C_6H_5C(NH)OCH_3.HCl$ . This salt gave, on being heated, methyl chloride and benzamide, and, on treatment with water, it gave methyl benzoate and ammonium chloride. Finally, treatment of the original chlorimide with hydrogen sulphide and distillation of the product with alkali failed to give even a trace of methylamine—which showed that the isomeric methyl derivative had not been formed even in traces.

The salts of the acid chloramides, accordingly, have the constitution (II), corresponding to this ether. We also found that benzchloramide could be titrated quantitatively with a standard alkali, using phenolphthalein as the indicator.

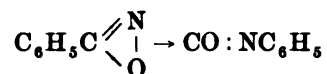
A little later both of these results were also obtained, independently, by Hantzsch; the latter showed by his elegant method of investigation that the free acid halogen amides are no true acids; that the metal derivatives, however, are the salts of true and not very weak acids, and consequently derivatives of a substance of different constitution from the free halogen amides. By analogy to certain corresponding ketones and nitro compounds, Hantzsch assigns to the salts the constitution (II), as salts of an unstable hydroxyl acid,  $RC(OH)(NHal)$ , to the neutral parent substance the constitution of an amide,  $RCO.NHAl$ . It is evident that the two methods of experimentation, the physico-chemical and the organo-synthetic, supplement each other in a desirable fashion.

This not unexpected result of the determination of the constitution of the salts of the acid bromamides does not affect the interpretation of the "Beckmann rearrangement" as given above; the change



is quite normal in view, notably, of Thiele's<sup>45</sup> work on dissociation and addition reactions involving two even more widely separated atoms in a molecule than are the nitrogen and oxygen atoms in the salts of the acid bromamides.

The constitution of the salts also suggests, however, a loss of potassium bromide and the formation of a ring compound,  $RC:NO$ , rather than a univalent nitrogen derivative as the intermediate product which actually suffers the molecular rearrangement. The strain in the triatomic ring, rather than the reactivity of univalent nitrogen, would then be the ultimate cause of the rearrangement:



Such a suggestion has been made, incidentally, by Hantzsch<sup>46</sup> in explanation of the rearrangement of the acetate of benzhydroxamic acid, and Baeyer<sup>47</sup> has expressed a similar view in connection with the transformation of ketone peroxides into lactones. It may be said here that such an interpretation could be made even far more general; as the compound,  $RCO.N:$ , could very readily go over into the isomeric ring derivative, many of the facts in favor of the formation of the former would obviously hold for the latter. The behavior of the acid bromamides, the hydroxamic acid salts and, possibly, of the acid azides, the contrasting behavior of the substituted acid halogen amides and hydroxamic acids,  $RCONC_6H_5(OH)$ , would be in equally good agreement with both interpretations. But the following facts favor, after all, decidedly, the more comprehensive idea of the agency of univalent nitrogen:

1. Amidoximes,  $RC(NO)NH_2$ , undergo the rearrangement<sup>48</sup> under the influence

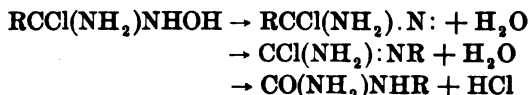
<sup>45</sup> *Liebig's Annalen*, Vol. CCCXCVI, p. 87.

<sup>46</sup> *Berichte d. deutsch. chem. Gesell.*, Vol. XXVII, p. 1256; see note 20, p. 8, for his most recent views.

<sup>47</sup> *Ibid.*, Vol. XXXII, p. 3627.

<sup>48</sup> *TIEMANN, ibid.*, Vol. XXIV, p. 4162.

of acid dehydrating reagents as do other oximes. An intermediate ring compound,  $\text{RC}=\text{N}-\text{NH}$ , a hydrazine or isodiazomethane, would tend to give hydrazine<sup>49</sup> or to change, if at all, into its closely related isomer, a diazo methane,  $\text{RCH}-\text{N}=\text{N}$ , and give the latter's decomposition products, nitrogen and an alkylidene derivative, rather than suffer a separation of two carbon atoms, from each other, and two nitrogen atoms, from each other, against their well-known tendencies. On the other hand, the reactions

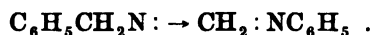


introduce no new assumption into the theory of the activity of univalent nitrogen—they appear as entirely normal.

2. The ring theory cannot be applied to the rearrangement of the ketoximes until an N-alkyl (aliphyl) aldoxime or ketoxime,  $\text{R}_2\text{C}:\text{NR}'-\text{O}$ , has been made to suffer a "Beckmann rearrangement" with a migration of an *alkyl* or an *aliphyl* group, R, since such a ring,  $\text{R}_2\text{C}-\text{NH}-\text{O}$ , would have to be the cause of the rearrangement of the ketoximes.<sup>50</sup> Such a migration does not appear on record, although the "Beckmann reagents" have been applied to such oximes.<sup>51</sup>

3. According to Werner,<sup>52</sup> the hydroximic acid chlorides,  $\text{RCCl}(\text{NOH})$ , and benzoates,  $\text{RC}(\text{O}_2\text{CC}_6\text{H}_5)(:\text{NOH})$ , in which the ready formation of the above ring appears particularly well favored, do lose acid easily, but no rearrangement ensues.

4. The ring theory is utterly incompatible with the most recent case of a "Beckmann rearrangement," that of benzyl azide,<sup>53</sup>  $\text{C}_6\text{H}_5\text{CH}_2\text{N}(\text{N}_2)$ , while the latter appears as a perfectly normal case of a rearrangement of a univalent nitrogen derivative, benzylimide:



Finally, attention is called once more to the fact that, whereas there is not a single fact concerning the constitution and reactions of organic compounds, as known to the author, which does not agree perfectly with the theory of the "Beckmann rearrangement" postulating the intermediate formation of a univalent nitrogen derivative, it does not agree so well with the more obscure relations of the theory of stereoisomerism of ketoximes and their influence on the rearrangement of these isomers.<sup>54</sup> It is hoped that future work will remove this difficulty.

The conclusions arrived at may be summed up as follows:

1. The interpretations of the "Beckmann rearrangement" given by Hoogewerff and Van Dorp, Hantzsch, Beckmann, Hesse, Freundler, and Nef do not agree with all the facts established in connection with this reaction.

<sup>49</sup> HANTZSCH, *ibid.*, Vol. XXXIV, p. 2507.

<sup>50</sup> *Vide* BAAYER, *loc. cit.*

<sup>51</sup> *E. g.*, BECKMANN, *ibid.*, Vol. XXVI, p. 2272.

<sup>52</sup> *Ibid.*, Vol. XXXII, pp. 1654, 1975; *vide* these papers for the actual products.

<sup>53</sup> CURTIUS, *loc. cit.*

<sup>54</sup> *Vide* the first paper, *loc. cit.*

2. All the most important results of the investigations of the reaction from the point of view of the constitution of the substances involved agree best with that interpretation which postulates the intermediate formation of a univalent nitrogen derivative as the essential cause of this characteristic rearrangement. This conception is comprehensive enough to satisfy all the cases where the rearrangement has been observed actually to occur in widely varied classes of compounds, such as:

I. Acid halogen amides:	$\text{RCONH}(\text{Hal})$
II. Acyl oximido acid salts:	$\text{RC}(\text{NOAcyl})(\text{OMe})$
III. Ketoximes:	$\text{R}_2\text{C}(\text{NOH})$
IV. Oximido acid esters:	$\text{RC}(\text{NOH})(\text{OR})$
V. Oximido acids:	$\text{RCO.NHOH}$
VI. Amidoximes:	$\text{RC}(\text{NOH})\text{NH}_2$
VII. Acid azides:	$\text{RCO.N}(\text{N}_2)$
VIII. Benzyl azides:	$\text{RCH}_2\text{N}(\text{N}_2)$

It also satisfies those otherwise analogous cases where the rearrangement for no other apparent reason than the impossibility of the fulfilment of this condition, could not be effected, such as

I'. Acyl alkyl (alphyl) halogen amines:	$\text{RCO}(\text{NHalR}')$
II'. Halogen imido esters:	$\text{RC}(\text{NHal})\text{OR}'$
III'. Acyl $\beta$ - alphylhydroxylamines:	$\text{RCO}(\text{NR}'\text{OH})$
IV'. N - Alkyl (alphyl) ketoximes:	$\text{R}_2\text{C}(\text{NR}'\text{O})$

3. Whatever interpretation be given of the rearrangement, it seems to have been satisfactorily established that the groups of compounds I and II suffer the transformation only when opportunity is offered for the loss of the acids, H (Hal) or HO (Acyl), by the formation of their salts, groups III, IV, V, VI when dehydrating reagents are used, VII when nitrogen is evolved under the influence of heat.



**DIMENSIONS OF DIRECT-CURRENT DYNAMO-  
ELECTRIC MACHINES**





# A SYSTEMATIC METHOD OF CALCULATING THE DIMENSIONS OF DIRECT-CURRENT DYNAMO-ELECTRIC MACHINES

CARL KINSLEY

THE many types of dynamos and motors used vary from the toy fan motor to the monster machine which furnishes power to a city, from the low-voltage large-current dynamo for electro-plating to the high-voltage constant-current machine for arc lighting. In all cases there are, however, certain underlying principles which must always be used in proportioning the parts of any new machine.

It is the purpose of this paper to arrange the equations showing the relations between the most important constants of any machine in such a manner that those details may readily be determined whenever the type of machine desired is known.

The conditions of each problem will determine the voltage, current, and speed needed. The rise of temperature that is allowable and the efficiency desired are the most important considerations for the designer. The total efficiency depends largely on the mechanical details and the care used in building the shaft and bearings of the machine.

In every case, then, the design will be made to give the rise in temperature and efficiency desired. All of the designing constants necessarily assumed must conform to good practice, but will unite to produce the predetermined results.

The rise in temperature now allowed in the best machines may be taken at 40° C. in the armature at full load and 35° C. in the fields. Ten years ago twice this value was permitted, but experience has shown that the above values are much preferable.

The electrical efficiency varies with the size of the machine under consideration. For a direct-current, constant potential generator the following would be fair values when the peripheral speed is about 1,700 cm. per second. A reduction of peripheral speed to 1,300 cm. per second will reduce the efficiency expected only 1 per cent.:

1,500 K.W. will have 99.00 per cent.	25 K.W. will have 96.00 per cent.
1,000 K.W. " " 98.70 per cent.	15 K.W. " " 95.00 per cent.
500 K.W. " " 98.50 per cent.	10 K.W. " " 94.50 per cent.
100 K.W. " " 98.00 per cent.	5 K.W. " " 93.50 per cent.
50 K.W. " " 97.50 per cent.	1 K.W. " " 91.50 per cent.

As an illustration suppose that the machine desired is as follows:

Direct-current machine giving a constant voltage of 115 volts.	
Current at full load - - - - -	260 amp.
Rated power - - - - -	30 K. W.
Speed - - - - -	950 r. p. m.

Assume a belted four-pole type, with parallel drum wound armature and compound wound field with long shunt connection.

Before discussing a new method of design, those in common use or recently proposed will be noted.

In the manufacturing companies of the United States the designing consists largely of ingrafting on the existing types of machines the improvements that experience suggests. The original machines are seen to have been exceedingly crude and faulty when compared with the present designs. The careful economy in the use of material also requires that one design shall depend somewhat upon the material left from the construction of others. The best machines on the market have finally become nearly perfect, from long years spent on tests and on the corrections of faults.

In order to be able intelligently to test machines and to suggest needed improvements, a consistent method of proportioning the parts of a machine must be employed. It is therefore necessary to have a thorough understanding of the fundamental relations of the dimensions of a machine to its performance.

It might be of interest to outline the method used by Wiener, in his recent and exhaustive book<sup>1</sup> on the subject, as it is a typical one. It is assumed that the type of machine, the electro-motive force, the current and the revolutions per minute are necessarily of certain values. Six assumptions are then made by the aid of tables showing the results of modern practice. The remaining calculations follow naturally from them.

*First assumption* (p. 52).—The conductor velocity is chosen, *i. e.*, the peripheral velocity of the armature,  $V_c$ . From 25 to 100 feet per second may be used, but 50 is recommended for the ordinary machine.

Since the revolutions per minute are known, the diameter,  $d_a$ , of the armature may be computed.

*Second assumption* (p. 53).—Field density of magnetization or average magnetic density in the air gap  $\mathfrak{S}_a$ , is now assumed.

Since the e. m. f. desired is known, the total length of the active wire on the armature,  $L_a$ , may be computed.

*Third assumption* (p. 56).—The carrying capacity of the armature wire, circular mils per ampere, is given a value conforming to good practice.

The size of the armature conductors,  $\delta_a^2$ , may then be found.

*Fourth assumption* (p. 66).—The relative size of slot,  $b_s$ , and tooth,  $b_t$ , on armature is now taken.

From the diameter of the armature, already computed, the number of slots,  $N'_c$ , may now be determined.

From the size of the wire chosen, the total number of armature conductors,  $N_c$ , immediately follows.

Now, from the total length of active conductors and their number, the length of the armature,  $l_a$ , may be obtained.

*Fifth assumption* (p. 54).—The magnetic density in the armature core,  $\mathfrak{B}_a$ , is

<sup>1</sup> *Practical Calculation of Dynamo-Electric Machines.*

given a value depending on good practice. It may vary with the type of machine from 4,500 to 20,000 lines per square centimeter.

As the total magnetic induction through the armature may have been already computed, the radial thickness of iron on the armature,  $b_a$ , follows.

*Sixth assumption* (p. 95).—Finally the ratio between the total length of wire on the armature and that which is active is estimated.

Having the size of the wire and the total length, the resistance of the armature,  $r_a$ , may now be obtained.

The energy loss in the armature is now calculated. If the efficiency is found to be an unsatisfactory value, the above calculations are again made with different assumptions.

An estimation of the temperature rise is now made, based on an empirical formula. If the value found is not satisfactory, the design is again made and other assumptions are chosen. This method of solution by trial and error is frequently very laborious for one who has not had much experience in designing. A number of methods have been proposed to simplify the process.

Monnier<sup>2</sup> recently proposed a method by which the relation between the diameter and the length of the armature could be determined upon the assumption of various designing conditions. He considers, however, only the joule loss in the copper conductors of the armature, and neglects the equally important hysteresis and eddy current losses. The very important consideration of the efficiency of the machine seems also to have been overlooked.

Pasqualini<sup>3</sup> gives an interesting discussion of the problem and develops a logical method of proportioning the armature of a dynamo. Unfortunately, however, he is in error in stating that "there is a single condition from which it is logical to start, namely, that of a minimum amount of iron in the armature; for it is this condition that assures that the machine will operate with a minimum loss from hysteresis and consequently with the highest efficiency." This is on the basis that the joule loss in the armature and the magnetic density of the core are fixed. Neither of these quantities is fixed, and there is no minimum loss by hystereses and eddy currents, as is shown by equations (11) and (13). Also, it is necessary to recalculate the dimensions, if after the completion of the design, it is found that not sufficient radiating surface has been provided.

In 1896 I first arranged a systematic method of proportioning the dimensions and winding of a dynamo. A great many designs were made at that time, but were not published. The method then used has recently been amplified, and the results of recent improvements, adopted by manufacturers, have been considered.

It is well to emphasize again the characteristics which will make the machine designed a successful one:

<sup>2</sup> *Traité d'électricité industrielle*; see JACKSON, *Electro-Magnetism*, p. 119.

<sup>3</sup> "Italian Association of Electrical Engineers," July, 1902. See *Electrical World*, October 18.

1. The maximum rise of temperature at full load should be equal to 40° C.

2. The electrical efficiency should be of the desired value, equal to 96.34 per cent.

This efficiency should be a maximum at the particular outside load most frequently used. The hysteresis and eddy current losses should also be controlled, so that the machine losses may not exceed the value desired.

It is not sufficient that these losses be made as small as possible, since the material used in construction increases as the loss decreases. In computing the dimensions of a machine, the electrical efficiency taken from the preceding table and the relative value of copper and iron losses, given on p. 8, may be used. The resulting dynamo will have the dimensions which modern practice has found to be entirely satisfactory.

The various resistances of the parts of the dynamo circuit, upon which the efficiency depends, may be computed as follows:

$$\eta_e = \frac{EI}{E'I'} = \frac{1}{1 + \frac{r_s}{R} + 2 \frac{r_s}{r_{sh}} + \frac{(r_s + r_{sh})R}{r_{sh}}} \quad (1)$$

$\eta_e$  = electrical efficiency.

$E$  = voltage on the terminals of the machine.

$E'$  = electro-motive force developed by the machine.

$I$  = current used outside of the machine.

$I'$  = total current through the armature conductors.

$R$  = resistance of the circuit outside of the dynamo.

$r_{sh}$  = resistance of the shunt windings.

$r_s = r_a + r_{se}$  = resistance of the windings carrying the current  $I'$ .

$r_a$  = resistance of the armature.

$r_{se}$  = resistance of the series windings on the field.

The maximum efficiency of the machine will occur at some particular load on the machine, depending upon the relative resistances. It is of advantage to make this maximum occur at some predetermined load.

Make

$$\frac{d\eta_e}{dR} = 0$$

and then solve for  $R$

$$R = r_{sh} \sqrt{\frac{r_s}{r_s + r_{sh}}} \quad (2)$$

Substituting this condition for maximum efficiency in equation (1),

$$\eta_e = \frac{1}{1 + 2 \frac{r_s}{r_{sh}} \sqrt{\frac{r_s + r_{sh}}{r_s}} + 2 \frac{r_s}{r_{sh}}} \quad (3)$$

Since  $r_s$  is very small, compared with  $r_{sh}$ ,  $r_s + r_{sh} = r_{sh}$  and  $\frac{r_s}{r_{sh}}$  may be neglected.

\* LORD KELVIN in *La lumière électr.*, Vol. IV (1881), p. 385, gives the equation for a shunt machine which can be extended to cover the case of the compound machine.

Equation (3) then becomes

$$\eta_e = \frac{1}{1 + 2\sqrt{\frac{r_s}{r_{sh}}}}.$$

Solve for  $\frac{r_{sh}}{r_s}$

$$\frac{r_{sh}}{r_s} = \left( \frac{2\eta_e}{1 - \eta_e} \right)^2. \quad (4)$$

Substitute in equation (4)  $\eta_e = 96.34$  per cent.

$$\frac{r_{sh}}{r_s} = 2765.$$

Substitute this in equation (2) and make  $r_s + r_{sh} = r_{sh}$ .

$$r_{sh} = \sqrt{2765} R = 52.6 R.$$

$$R = \frac{E}{I} = \frac{115}{260} = 0.443 \text{ ohms.}$$

$$r_{sh} = 23.3.$$

$$r_s = \frac{2765}{r_{sh}} = 0.00844.$$

In this type of machine it is safe to take  $r_a = 0.85 r_s$ ; then

$$r_a = 0.00717,$$

although in over-compounded machines the value of  $r_a$  may sometimes be only one-half of the value of the above, which is also exclusive of the brush resistance.

These values for resistance will thus give an electrical efficiency, maximum at full load, of 96.3 per cent. At one-half of full load the efficiency will be 95.5 per cent.

Since a multiple winding is to be used on the armature, there will be four paths through it.

Then

$$R_a = 16 r_a = 0.1146.$$

$$L_t = \frac{R_a}{0.942 \times 10^{-8}} = 12170 \text{ cm.} \quad (5)$$

$$L_a = 0.4 L_t = 4870 \text{ cm.}$$

$R_a$  = resistance of all the armature conductors in series.

$L_t$  = the total length of the conductors.

$L_a$  = the length on armature facing the field poles.

$0.942 \times 10^{-8}$  = resistance of 1 cm. of No. 6 B.W.G.

This size of wire gives 600 circular mils per ampere, which has been found to be a safe value for armatures of this type.<sup>5</sup>

<sup>5</sup> WIENER, *Dynamo Electric Machine*, 2d ed., p. 39.

<sup>6</sup> *Ibid.*, p. 56.

The value  $L_a = 0.4 L_t$  is that used in some machines by good makers. A value of  $L_a = 0.3 L_t$  would be more nearly the average, however.

One of the determining equations can now be written. It is dependent upon the assumed efficiency.

$$N_c \frac{d_a}{k_1} = 4870 \quad (A)$$

$N_c$  = number of conductors on the armature.

$L_a$  = length of each conductor, or the length of the armature that faces the field poles.

$k_1$  = relation between diameter and length of the armature.

$d_a$  = diameter of the iron in the armature.

Equation (A) has thus been obtained by assuming a desired efficiency for the dynamo.

The rise in temperature of the armature depends upon (1) the heat dissipated in it, (2) the radiating surface exposed, and (3) its peripheral speed.

At a peripheral speed of 3,500 feet per minute (which is about what the high-speed machine under consideration will give), and an open type of armature windings (such as is now used by the leading manufacturers), about 2.2 watts per square inch (0.341 watts per square centimeter) of the armature core will give the desired rise of 40° C. in the armature.<sup>7</sup>

The hysteresis and eddy current losses can be reduced to a low value by increasing the iron and so decreasing the magnetic induction of the core of the armature, as is shown by equations (8), (11) and (13).

A value conforming to good practice may be taken as follows:

$$P_A = P_h + P_e + P_a = 2.2 P_a.$$

$$P_a = (I')^2 R_a = 502.5, \therefore P_A = 1105 \text{ watts.}$$

$P_A$  = total energy dissipated in the armature.

$P_h$  = energy lost by hysteresis.

$P_e$  = energy lost by eddy currents.

$P_a$  = energy lost by heating of copper.

$I_{sh}$  = current through shunt windings on field.

$$I_{sh} = \frac{E}{r_{sh}} = \frac{115}{23.3} = 4.94 \text{ amp.}$$

$I' = I + I_{sh}$  = current in armature conductors, 265 amp.

$R_a$  = resistance of armature = 0.00717.

This constant, depending as it does on the number of reversals per second, will decrease with a decrease in the velocity. It may be made to follow closely the change in the radiation constant.

The radiation per square centimeter of radiating surface of the armature will depend upon the type of armature and peripheral velocity. If a high-speed machine

<sup>7</sup> JACKSON, *Electro-Magnetism*, p. 106.

is considered, having an exposed type of winding, the following relation will be true:

$$\frac{P_A}{S_A} = \frac{\theta_a}{320} \left( 1 + \frac{V}{1000} \right) .$$

$\frac{P_A}{S_A}$  = watts per square centimeter to be dissipated by the armature.

$V$  = peripheral velocity in centimeters per second.

$\theta_a$  = rise of temperature in degrees C.

If  $\theta_a$  is 40° C., and  $V = 1735$  centimeters per second, as is the case of the high-speed Westinghouse machine of this output, then

$$\frac{P_A}{S_A} = 0.341 .$$

The slow-speed machine (550 r. v. m.) of this same output, however, has a constant

$$\frac{P_A}{S_A} = 0.27 .$$

Take, therefore, a constant depending upon the speed of the machine to be designed.

The equation depending upon the rise of temperature in the armature can now be written:

$$S_A = \pi d_a l_a + \frac{\pi d_a^2}{4} = d_a^2 \left( \frac{\pi}{k_1} + \frac{\pi}{4} \right) .$$

Take only one-half of each end on account of the smaller average velocity.

This must be equal to

$$\frac{A}{0.341} = 3242 \text{ sq. cm.} = S_A . \quad (6)$$

$$d_a^2 \left( \frac{\pi}{k_1} + \frac{\pi}{4} \right) = 3242 . \quad (B)$$

$S_A$  = surface of the armature core. This may be taken as a fair approximation, although the core is a hollow instead of a solid cylinder. The constant of 0.341 watts per square centimeter would have to be changed if the surface were computed in another way.

$d_a$  = diameter of armature.

$l_a$  = length of armature.

$$k_1 = \frac{d_a}{l_a} .$$

Let us now consider the electro-motive force of this machine. Since the e. m. f. depends merely on the number of lines cut per second, the equation can readily be



## 10 DIMENSIONS OF DIRECT-CURRENT DYNAMO-ELECTRIC MACHINES

written. Assume the constants which are entirely satisfactory from the standpoint of good practice.

$$\beta_1 = 0.70. \quad \mathfrak{S}_A = 7800 \text{ lines per square centimeter.} \quad (7)$$

$$\phi' = \frac{\pi d_a l_a \beta_1}{4} \mathfrak{S}_A.$$

$$E = \frac{4 \phi' n' N_c}{4 \times 10^8} = \frac{\pi d_a^2 \beta_1 n' N_c}{k_1 \times 4 \times 10^8} \mathfrak{S}_A.$$

Dividing by  $10^8$  gives the e. m. f. in volts. The constants 4 in this equation are due to the connections which provide that there are four parallel circuits through the armature, and that each conductor cuts  $\phi'$  lines four times during each revolution.

$$\frac{N_c d_a^2}{k_1} = \frac{E \times 4 \times 10^8}{\pi \beta_1 \mathfrak{S}_A \cdot n'} = \frac{115 \times 4 \times 10^8}{\pi \times 0.7 \times 7800 \times 15.83} = 169500. \quad (C)$$

$$\frac{N_c d_a^2}{k_1} = 169500.$$

$\beta_1$  = part of the armature surface covered by the poles of the field.

$\mathfrak{S}_A$  = magnetic density in the air gap.

$\phi'$  = number of lines coming from one pole.

$d_a$  = diameter of the armature.

$l_a$  = length of the armature.

$n'$  = revolutions per second.

$N_c$  = number of conductors on the armature.

The three equations (A), (B), and (C) were obtained by considering the efficiency, the rise in temperature, and the electro-motive force of the dynamo. The solution of these equations will give the necessary dimensions of the machine. The intermediate assumptions control these dimensions, but whatever are the assumptions made, the resulting dimensions obtained by solving these equations will give a machine which will operate in the predetermined manner. If for any reason the dimensions are undesirable, a change in an intermediate assumption may be made—such as  $\mathfrak{S}_A$ , for instance—and the three equations solved again. As long as the design constants are taken in accordance with good practice, the resulting dimensions will give a machine which the experience of various manufacturers has shown to be satisfactory.

The assumptions chosen are such as will produce a machine similar to the Westinghouse 30 K.W., direct current belted generator at 950 r. p. m.

These design constants are as follows:

$$\begin{array}{ll} L_a = 0.4 L t^8 & \text{equation (5)} \\ P_a = 0.341 S_A^9 & \text{equation (6)} \\ \beta_1 = 0.70^{10} & \text{equation (7)} \end{array} \quad \begin{array}{ll} \mathfrak{S}_A = 7800^{11} & \text{equation (7)} \\ \delta_a^2 = 600 \text{ circular mils per amp.}^{12} & \end{array}$$

<sup>8</sup> WIENER, *Dynamo-Electric Machines*, 2d ed., p. 100. The constant used here was from a particular Westinghouse machine.

<sup>9</sup> JACKSON, *Electro-Magnetism*, p. 106. In the type of armature adopted, this will give a rise in temperature of about 40° C

<sup>10</sup> A common value and the one used in the Westinghouse machine.

<sup>11</sup> A usual value which will give an economical value of  $\mathfrak{S}$  in the armature teeth.

<sup>12</sup> WIENER, p. 56. A safe value for imbedded conductors.

The three equations are

$$(A) \quad \frac{N_c d_a}{k_1} = 4870 .$$

$$(B) \quad d_a^2 \left( \frac{\pi}{k_1} + \frac{\pi}{4} \right) = 3242 .$$

$$(C) \quad \frac{N_c d_a^2}{k_1} = 169500 .$$

From these equations the dimensions are found to be:

Diameter of the armature  $= d_a = 34.9$  cm.

Length of the armature core  $= l_a = 20.85$ .

Number of conductors on the armature  $= N_c = 234$ . Use 236 for winding reasons.

The size and number of teeth on the armature can now be determined. The depth of the slot should be three or four times the width. In this case take four wires in a slot of one wire to a layer. The width of the slot,  $b_s$ , to take the wire, its insulation, and the insulation on the sides of the slot will be equal to 0.70 cm.

The width of a tooth at the surface  $= b_t$ .

$$b_t = \frac{\pi d_a}{\frac{1}{4} N_c} - b_s = 1.86 - 0.70 = 1.16 \text{ cm.}$$

The magnetic density in the iron at the surface will be

$$7800 \times \frac{1.86}{1.16} = 12500 \text{ lines per square centimeter.}$$

The length of the air gap is of great importance and it should be determined to make sure that this is a possible design.

When the armature carries current, it is changed into a magnet which at the weak pole corners acts against the fields. The magneto-motive force (measured in ampere turns) due to the fields must always be stronger than that due to the armature, in order to prevent a reversal of the field and an accompanying shift of the neutral diameter. Should this occur it would lead to destructive sparking in a machine of this type, where the brushes are not automatically moved.<sup>13</sup> Make therefore the ampere turns on the fields required for the air gaps,  $at_g$ , 12 per cent. more than the armature ampere turns.

$$at_g = 1.12 \times \frac{I'}{4} \times \frac{N_c}{4} = 4370 .$$

Since

$$\mathfrak{F}_A = \frac{4\pi at_g}{10 l_g} ,$$

$$l_g = \frac{1.26 at_g}{\mathfrak{F}_A} = \frac{1.26 \times 4370}{7800} = 0.706 \text{ cm.}$$

$l_g$  = length of double air gap. Consequently the clearance between the poles and the armature is 0.35 cm

<sup>13</sup> H. J. RYAN, *Trans. Am. I. E. E.*, Vol. VIII, p. 455



If the machine is to be operated at more than full load, the air gap could be advantageously lengthened.

In order to keep the e. m. f. at the terminals of the machine constant, there should be placed on the field series windings which would have the effect of neutralizing (1) the lengthening of the air gap due to the cross magnetization of the armature, (2) the demagnetizing effect of the armature ampere turns within the double angle of lead of the brushes, and (3) the fall of potential through the armature and series windings when the conductors are carrying a current.<sup>14</sup>

The dimensions determined in the above design are substantially the same as those found in a Westinghouse dynamo of the same type:

Westinghouse dynamo.

$d_a = 35.0$  cm.     $l_a = 20.2$  cm.

Armature clearance = 0.35.

$N_c^1$  = number of slots = 61.

$b_t = 1.16$ .     $b_s = 0.70$ . Although a copper bar is used instead of a wire.

Finally, it will be of interest to determine the hysteresis and eddy current losses in the armature, as a function of the magnetic induction in the iron core.

The total number of lines passing into the core from one pole of the field is known.

$$\phi' = \frac{\pi d_a l_a}{4} = 3145000 .$$

As these lines pass through the core after traversing the teeth equally in each direction,

$$\begin{aligned} b_a l_a \mathfrak{B}_a &= \frac{\phi'}{2} . \\ b_a &= \frac{75300}{\mathfrak{B}_a} . \end{aligned} \tag{8}$$

$b_a$  = radial thickness of the iron on the armature.

$l_a$  = length of the armature core.

$\mathfrak{B}_a$  = magnetic induction in the iron of the core.

The iron losses can now be computed.

$$P_A = 10^{-7} 0.0035 \mathfrak{B}_a^{1.6} n' M_1' \text{ in watts} . \tag{9}$$

$P_A$  = hysteresis loss.

$n_1$  = periods per second =  $n' \times n_p = 15.83 \times 2 = 31.66$ .

$n_p$  = number of poles = 4.

$k'$  = part of volume filled with iron. For iron sheet,  $k' = 0.90$ .

$M_1'$  = volume of iron in core.

$$= b_a l_a k' \times \pi (d_a' - b_a) .$$

$d_a'$  = diameter at bottom of teeth.

<sup>14</sup> See JACOBSON, *Electro-Magnetism*, pp. 158, 218.

Since

$$l_a = 20.85, \quad d_a = 34.9, \quad d'_a = 28.7 \text{ cm.}, \quad \text{and} \quad k' = 0.90.$$

$$M'_1 = 1690 b_a - 5.90 b_a^2. \quad (10)$$

Substitute equation (8) in equation (10):

$$M'_1 = \frac{12.7 \times 10^7}{\mathfrak{B}_a} - \frac{3.34 \times 10^{11}}{\mathfrak{B}_a^2}.$$

Substitute in equation (9):

$$P_h = 1.405 \mathfrak{B}_a^{2.6} - 3700 \mathfrak{B}_a^{-4.4}. \quad (11)$$

This equation was obtained from experimental data by Steinmetz.<sup>15</sup>

The exponent of  $\mathfrak{B}$  has been questioned by Ewing and Kleassen,<sup>16</sup> but the error was found to be small, even over a much wider range of values of the magnetic induction than those used in practice. Recently, Doery<sup>17</sup> has attempted a determination of the equation from theoretical instead of experimental considerations. He finds that the equation is substantially correct, the exponent of  $\mathfrak{B}$  very properly being 1.6 as the above experimenters have found.

The eddy current losses depend largely upon the thickness of the iron plates. This loss is not large under normal conditions.

$$P_e = 1.645 \times 10^{-11} n_1^2 M_1' \mathfrak{B}_a^2 \delta^3. \quad (12)$$

$P_e$  = eddy current loss in watts.

$\delta$  = thickness of iron plates in centimeters = 0.05.

Substitute values found above for  $M_1'$  and  $n_1$ :

$$P_e = 0.0524 \mathfrak{B}_a - 13.75.$$

The hysteresis loss in the teeth is constant, whatever the induction in the core may be. From equation (11) and equation (13) it is seen that there is no value of  $\mathfrak{B}_a$  and consequently of  $b_a$  which will make the iron loss a minimum.

The value for  $\mathfrak{B}_a$  which will make  $P_h + P_e = 1.2 P_a = 604$  watts, as assumed, can readily be determined.

If the depth of the slot is taken at 3.1 cm., which will easily accommodate the four wires, the volume of iron in the teeth is 3,880 cubic centimeters. The average magnetic induction is 14,500. Substitute these values in equation (11) and the hysteresis loss is then found to be 196.3 watts. Substituting the same values for  $\mathfrak{B}_a$  and  $M_1'$  in equation (12), the eddy current loss in the teeth is found as 33.7 watts. Total iron loss in the teeth equals 230 watts. If the induction in the core is taken at 10,000 lines per square centimeter, substituting in equation (11) gives the hysteresis

<sup>15</sup> STEINMETZ, *Alternating Current Phenomena*, p. 116.

<sup>16</sup> *Phil. Trans. Roy. Soc.*, Vol. CLXXXIV (1893), p. 1017.

<sup>17</sup> *Zeit. f. Elek.*, July 6, 1902.

<sup>18</sup> STEINMETZ, p. 135. This equation is derived from theoretical considerations. It has been universally em-

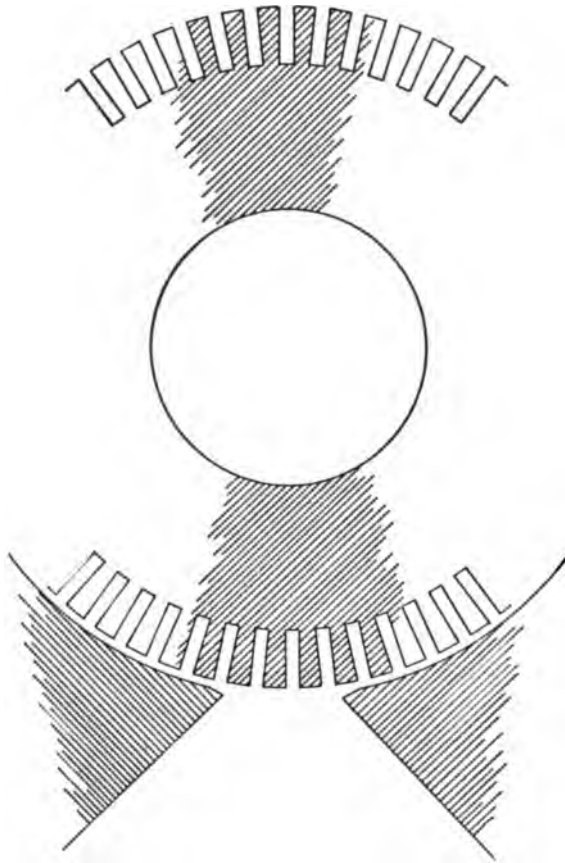
ployed for many years in the calculation of eddy current losses. THOMSON, *Trans. Am. I. E. E.*, April, 1902, has found experimentally that the exponent of  $\mathfrak{B}$  should be 1.6 instead of 2. There exists also some logical ground for this result, but as the eddy current loss is small the old equation will be used in these calculations.

loss as 262 watts. Substituting again in equation (13) the eddy current loss is found to be 39 watts.

Total iron loss in core = 301 watts.

Total iron loss in armature = 531 watts.

This value is sufficiently near that assumed, so that an induction in the core of 10,000 lines per square centimeter may be adopted. Substitute this in equation (8) and the radial depth of the core is found to be 7.53 cm.



The accompanying drawing gives the proportions of the armature. It shows also the relation between the length of the air gap and the distance between the pole corners.

The rest of the calculations, including the size of the frame, the details of the commutator, and the completion of the calculation of the field windings, can be readily made in a manner suggested by any book giving modern practice.<sup>19</sup> The resistance of the field, both shunt and series, has been given above, and the ampere turns needed to drive the necessary magnetic induction through the air gap has already been calculated.

An important use for the equations given above might be pointed out. The makers protect, as professional secrets, a number of the design constants given above. They may be readily determined from any machine by substituting in the above equations the dimensions of the particular machine examined and

then solving them anew for the design constants desired.

<sup>19</sup> See JACKSON, *Electro-Magnetism*, chap. 5.

# **EQUILIBRIUM IN AQUEOUS SOLUTIONS OF CARBONATES**



# EQUILIBRIUM IN THE SYSTEM COMPOSED OF SODIUM CARBONATE, SODIUM BICARBONATE, CARBON DIOXIDE, AND WATER

HERBERT N. MCCOY

THE conversion of dry sodium bicarbonate into the normal carbonate, water, and carbon dioxide requires but a very moderate degree of heat. Complete decomposition takes place rapidly at 100–110°.<sup>1</sup> Gautier found that no loss of weight occurs when the salt is placed in a vacuum at 25°. At 35° an appreciable loss had occurred at the end of four days.

Solutions of sodium bicarbonate are far less stable than the dry salt. Rose<sup>2</sup> observed that a solution of sodium bicarbonate loses carbon dioxide by evaporation and leaves nearly pure normal carbonate. The same change was found to occur, at ordinary temperature, when the solution was allowed to evaporate under such conditions that the carbon dioxide given off could be absorbed by caustic potash. Magnus<sup>3</sup> found that carbon dioxide is removed from a solution of sodium bicarbonate at ordinary temperatures by a stream of hydrogen; and Marchand<sup>4</sup> obtained a similar result by drawing air through a saturated solution of sodium bicarbonate at a temperature of 0°. At 38° Marchand found that about 95 per cent. of the bicarbonate was converted into normal carbonate, after drawing a rather large volume of air through the solution, and concluded that the change would be complete after sufficiently prolonged action of the air.

Dibbitts<sup>5</sup> has pointed out the errors in the determination of the solubility of bicarbonates arising from a failure on the part of previous experimenters to prevent loss of carbon dioxide. In his own work on such determinations the loss of carbon dioxide was reduced to a minimum by working in closed vessels which were nearly filled by the solutions. In connection with this work Dibbitts determined roughly the pressure with which carbon dioxide is evolved by saturated solutions of sodium and potassium bicarbonates at several temperatures. He obtained the following results with sodium bicarbonate:

TABLE I

Temperature	Pressure, in mm. of Mercury
15	120
30	212
40	356
50	563

<sup>1</sup> GAUTIER, *Berichte der deutschen chemischen Gesellschaft*, Vol. IX (1876), p. 1434.

<sup>2</sup> *Poggendorff's Annalen*, Vol. XXXIV (1835), p. 158.

<sup>3</sup> *Ibid.*, Vol. XL (1837), p. 590.

<sup>4</sup> *Journal für praktische Chemie*, Vol. XXXV ('815), p. 389.

<sup>5</sup> *Ibid.* (2), Vol. X (1874), p. 417.

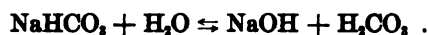


Treadwell and Reuter,<sup>6</sup> in connection with another problem, made a few determinations of the composition of a sodium bicarbonate solution of decinormal strength, when the partial pressure of the carbon dioxide was varied between 2 and 4 per cent. of one atmosphere. They concluded that the bicarbonate lost carbon dioxide at reduced pressures, but made no attempt to study equilibrium conditions.

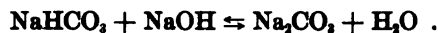
The consequences of a loss of carbon dioxide by solutions of sodium bicarbonate, in connection with analyses of carbonates, bicarbonates, and gases containing carbon dioxide, have at times been understood and taken into account. Thus Lunge<sup>7</sup> takes precaution to prevent loss of carbon dioxide from mixtures of carbonates and bicarbonates previous to and in the course of analysis. Lunge and Zeckendorf,<sup>8</sup> and also Rosenthal,<sup>9</sup> recognize the dissociation of sodium bicarbonate in aqueous solution as the cause of the incomplete absorption of traces of carbon dioxide by solutions of sodium carbonate. On the other hand, the change, during analysis, of sodium bicarbonate into normal carbonate was observed by Cameron<sup>10</sup> without, at the time,<sup>11</sup> the true cause being discovered.

The loss of carbon dioxide by solutions of bicarbonates is the necessary consequence of four things: (1) the weakness of carbonic acid as an acid; (2) the acid nature of sodium bicarbonate; (3) the ready dissociation of carbonic acid into water and carbon dioxide; (4) the comparatively small solubility of carbon dioxide in water.

The hydrolytic dissociation of sodium bicarbonate is represented by the equation:



The carbonic acid existing thus in solution will necessitate a certain gaseous pressure of carbonic dioxide, the ratio of the concentrations in the two phases being simply the solubility, as Ostwald uses the term. In a system composed of dilute sodium bicarbonate and vapor, the concentration of the carbon dioxide in the vapor must therefore be a direct measure of that of the free carbonic acid in the solution. It might seem then, at first thought, that by measuring the concentration of the carbon dioxide in the vapor in equilibrium with a solution of sodium bicarbonate of known strength one would be able to determine the degree of hydrolytic dissociation of the salt.<sup>12</sup> But the equilibrium in the system in question must be more complex than is indicated by the above equation, as the following considerations show: The immediate result of the action of water on the bicarbonate is doubtless the formation of some sodium hydroxide and free carbonic acid. This action would naturally be expected to reach a state of equilibrium by reason of the reverse reaction. However, the bicarbonate is itself an acid and can act upon the free sodium hydroxide forming water and the normal carbonate, thus:



<sup>6</sup> *Zeitschrift für anorganische Chemie*, Vol. XVII (1896), p. 202.

<sup>7</sup> *Zeitschrift für analytische Chemie*, Vol. XXXVII (1896), p. 689.

<sup>8</sup> *Ibid.*, Vol. XXXVI (1897), p. 329.

<sup>9</sup> *Ibid.*, p. 333.

<sup>10</sup> *American Chemical Journal*, Vol. XXIII (1900), p. 473.

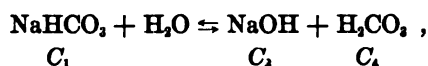
<sup>11</sup> The more recent work of CAMERON, *Journal of Physical Chemistry*, Vol. V, p. 537, is discussed later.

<sup>12</sup> See below, p. 11.

This last equation is really that of the hydrolytic dissociation of sodium carbonate. Though carbonic acid is much stronger, as an acid, than sodium bicarbonate, the relatively great concentration of the latter as compared with the former makes it necessary to consider the quantitative effect of the action of this acid-salt on the sodium hydroxide produced by the primary hydrolysis.

The equilibrium in aqueous solutions of sodium bicarbonate is therefore governed by the two reactions just discussed, each of which represents the hydrolytic dissociation of a salt of a weak acid and a strong base. Walker has shown<sup>13</sup> that for dilute solutions the product of the concentrations of base and acid divided by that of the salt is a constant. It is well known that the same relationship governs the hydrolytic dissociation of salts of strong bases and weak acids,<sup>14</sup> it being presupposed that both salt, and, in the latter case, base, are highly ionized and to practically the same extent.<sup>15</sup>

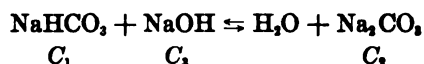
From the reaction



where  $C_1$ , etc., represent concentrations in gram-molecules per liter, we have

$$C_2 C_4 = K_1 C_1; \quad (1)$$

and from



we have

$$C_1 C_3 = K_2 C_5; \quad (2)$$

$K_1$  and  $K_2$  being the dissociation constants. From (1) and (2)

$$\frac{C_1^2}{C_2 C_4} = \frac{K_2}{K_1} = K. \quad (3)$$

$K$  is the equilibrium constant for the system under consideration.

In a paper entitled "The Solubility of the Alkali-Earth Carbonates in Water Containing Carbon Dioxide,"<sup>16</sup> Bodländer has worked out, in a different way from that given above, the equation governing the equilibrium in the reaction,



where  $\text{HCO}_3'$  and  $\text{CO}_3'$  refer to ionic concentrations. He gives

$$(\text{HCO}_3')^2 = \text{CO}_3' \times \text{CO}_2 \times \text{a constant}.$$

This equation and (3) are practically identical for dilute solutions, where one may neglect the factors for ionization. Bodländer applied this equation to the experiments of Schloesing<sup>17</sup> on the solubility of the carbonates of calcium and barium in water containing carbon dioxide, and to those of a similar nature of Engel<sup>18</sup> on magnesium

<sup>13</sup> *Zeitschrift für physikalische Chemie*, Vol. IV (1889), p. 324.

<sup>16</sup> *Zeitschr. f. physik. Chemie*, Vol. XXXV (1900), p. 32.

<sup>14</sup> VAN'T HOFF, *Vorlesungen über theoretische und physikalische Chemie* (Braunschweig, 1901), Vol. I, p. 124.

<sup>17</sup> *Compt. rend.*, Vol. LXXIV (1872), p. 1552; Vol. LXXV (1872), p. 70.

<sup>15</sup> ARRHENIUS, *Zeitschr. f. physik. Chemie*, Vol. V (1890), p. 17.

<sup>18</sup> *Ibid.*, Vol. C (1885), pp. 144, 352.

carbonate. Good results were obtained for the equilibrium constant. In the last paragraph of this paper Bodlander shows that the cause of increased solubility of sodium bicarbonate in water, when shaken with air, is due to loss of carbon dioxide and consequent formation of sodium carbonate. The calculated increase of solubility agrees approximately with the observed value. The discussion regarding the carbonates of sodium is brief, and no reference is made to other work on this subject.

In equation (3)  $C_4$  may be expressed in terms of the partial pressure,  $P$ , of the carbon dioxide and a constant,  $k$ , whose value depends upon the solubility of carbon dioxide in water, thus:

$$C_4 = kP .$$

According to Just<sup>19</sup> one liter of water at 25° dissolves 0.8256 liter of carbon dioxide measured at 25° and 760 mm. The concentration of such a solution is, therefore, 0.0338 normal molecular.

$$C_4 = 0.0338P .$$

$C_1$  and  $C_2$  may readily be expressed in terms of a single variable. Let one gram-molecule of sodium bicarbonate be dissolved in the volume  $V$ , of water. Let  $x$  represent the number of gram-molecules of sodium bicarbonate after equilibrium has been reached.  $1-x$  is the number of gram-molecules of bicarbonate changed into carbonate, and  $1/2 (1-x)$  is the number of gram-molecules of carbonate formed.

As is shown on p. 13, the concentration of the hydroxide formed is so minute that it may be left out of consideration in this connection. Its maximum amount did not exceed 0.2 per cent. of either  $x$  or  $1-x$  for the solutions studied.

$$C_1 = \frac{x}{V} . \qquad C_2 = \frac{\frac{1}{2}(1-x)}{V} .$$

Therefore

$$\frac{C_1^2}{C_2 C_4} = \frac{2x^2}{0.0338 P V (1-x)} = K .$$

Calling the concentration of the sodium in gram-atoms per liter  $C$

$$C = \frac{1}{V} ,$$

the equilibrium equation then becomes

$$\frac{2x^2 C}{0.0338 P (1-x)} = K .$$

The present work was undertaken with the view of testing the correctness of these deductions. After the experiments had been started a paper appeared<sup>20</sup> by Cameron and Briggs under the title "Equilibrium between Carbonates and Bicarbonates in Aqueous Solution." The authors drew air through aqueous solutions of the bicarbonates of sodium, potassium, and magnesium of various concentrations, and at four

<sup>19</sup> *Zeitschr. f. physik. Chemie*, Vol. XXXVII (1901), p. 342.

<sup>20</sup> *Journal of Physical Chemistry*, Vol. V (1901), p. 537.

different temperatures, and determined by analysis the proportions of carbonate and bicarbonate after equilibrium had been reached. It is rather remarkable that no reference is made, in their paper, to any previous record of loss of carbon dioxide by bicarbonate solutions. Inasmuch as no quantitative application of the law of mass-action to this equilibrium was attempted by Cameron and Briggs, the work here reported was continued.

The experimental problem presented was the determination of the composition of the liquid and vapor phases in equilibrium with one another at constant temperature, for various proportions of the constituents. The temperature chosen was 25°. The method adopted consisted in shaking 200 or 300 c.c. of the solution of the mixed carbonates contained in a closed liter bottle with air until equilibrium was reached, the whole being kept at constant temperature. A sample of the gas was transferred to another vessel and its percentage of carbon dioxide determined. The composition of the solution was determined by analysis.

A number of the best methods of analyzing carbonate solutions have been critically studied by Küster.<sup>21</sup> The result of this careful piece of work was to show that, of all methods, that of Winkler is most accurate. This method, as applied to solutions containing bicarbonates, consists in adding, in excess, a known volume of standard alkali, free from carbonate, precipitating the carbonate by addition of barium chloride, and finally titrating the excess of alkali with hydrochloric acid, using phenolphthalein as indicator.

This method was tested by means of preliminary analyses, carried out as follows: The necessary amount of barium chloride solution was first brought into an 80 cc. flask, then the measured volume of standard alkali was added, and finally 10 cc. of the carbonate solution was run into this directly from the pipette. When two or three analyses of the same solution were to be made the flasks were all charged with the barium chloride and standard alkali, and corked. Then the carbonate solution was measured from the pipette into each flask with as little loss of time as possible. After mixing, the flasks were tightly closed with rubber stoppers and allowed to stand a few minutes, after which the excess of alkali was titrated with decinormal hydrochloric acid. The net amount of alkali used (that is, the whole amount added less that neutralized by acid in the final titration) is a measure of the carbon dioxide in excess of that existing as normal carbonate.

In three determinations, using in each 9.90 c.c. of a bicarbonate solution, there were required, in each analysis, exactly 10.60 c.c. of the standard barium hydroxide solution. In three other determinations, using in each 9.90 c.c. of a different bicarbonate solution, there were required 10.42, 10.42, and 10.40 c.c. of barium hydroxide. By working in the manner just described, no carbon dioxide could be lost by the bicarbonate solution in the course of the analysis, and practically none absorbed from the air by the standard alkali.

<sup>21</sup> *Zeitschr. f. anorg. Chemie*, Vol. XIII (1897), p. 127.

Carbon dioxide may readily be lost during analysis from solutions rich in bicarbonate, if care be not taken to render them alkaline at once. Low results were invariably obtained when, instead of working in the manner described, the bicarbonate solutions were measured into beakers and allowed to stand, even for four or five minutes, before adding an excess of alkali.

The concordant results given above, together with those of the elaborate work of Küster, are sufficient to show the accuracy of the analytical method used in the determination of the carbon dioxide in the solutions.

In order to calculate the amounts of carbonate and bicarbonate in any solution containing a known amount of carbon dioxide, it is necessary to know the total amount of sodium. This was determined by titration with standard hydrochloric acid, using methyl orange as indicator and taking for the end reaction the "normal color" produced by the indicator when dissolved in water saturated with carbon dioxide. This method, recommended by Küster,<sup>22</sup> gives excellent results.

For the determination of carbon dioxide in mixtures of this gas with air the sample was brought into a half-liter separatory funnel closed by a perforated stopper. An excess of barium hydroxide was run in rapidly from a burette by inserting the tip of the burette in place of the glass plug of the rubber stopper. A solid stopper was then substituted for the perforated one. In doing this a quantity of gas escaped equal in volume to that of the barium hydroxide added. The solution was shaken vigorously with the gas for ten minutes. (Absorption of the carbon dioxide seemed to be complete in from one to two minutes.) The excess of alkali was then determined by titration with standard hydrochloric acid. Phenolphthalein was used as indicator.

The transference of the gas from the equilibrium bottle to the vessel in which it was analyzed required some care; for, since the equilibrium depends on the temperature and pressure, any change in these conditions during the process of separation of the phases is not permissible. The equilibrium vessel consisted of a liter bottle. It was fitted with a rubber stopper through one hole of which passed the stem of a 100 c.c. separatory funnel, through the other hole a bent tube with glass stop-cock. The stem of the separatory nearly reached the bottom of the bottle. The carbonate solution, to the amount of 200 to 300 c.c., was placed in the bottle; the stopper, carrying the separatory funnel and tube, was inserted and carefully wired. The bottles thus charged were placed in a large thermostat; the temperature was kept constant at 25° ( $\pm 0.01^\circ$ ). While in the thermostat, the bottles were constantly shaken by means of a device giving a rather rapid rocking motion. The total pressure of the gas phase was made nearly equal to the normal atmospheric pressure by opening the stop-cock of the bent tube occasionally during the first hour or two after starting. The attainment of equilibrium under these conditions required two or three days, as shown by special experiments made to determine this point. In the recorded experiments the time allowed was five days or more. Immediately before removing a bottle from the ther-

<sup>22</sup> *Loc. cit.*

mostat, preparatory to the analysis of its contents, the gas pressure was determined by measuring, with a simple manometer, the excess or deficiency of pressure with respect to the existing barometric pressure. This was readily done by attaching the manometer to the bent tube having the stop-cock. The transference, for analysis, of the larger part of the gas to a large separatory of known volume was accomplished at constant (atmospheric) pressure in the following manner: The end of the bent tube of the equilibrium apparatus was connected by a short rubber tube with one tube of a three-way stop-cock. A second tube of the three-way cock was connected by a rubber tube with a perforated rubber stopper which fitted the large separatory into which the gas was to be transferred. The third tube of the three-way cock was attached to a thin-walled rubber bulb of about 50 cc. capacity. To transfer the gas, the stem of the small separatory of the equilibrium apparatus was filled with the carbonate solution by applying suction at the mouth of the separatory and opening its stop-cock, the other cock of the apparatus being meanwhile kept closed. The small separatory was then filled with pure dry mercury, and the rubber bulb and the tubes leading to the large separatory were rinsed with the gas which was displaced by allowing the mercury to run into the equilibrium bottle. The large separatory was completely filled with mercury and connected with the delivery tube from the three-way cock. It was supported above the equilibrium bottle in such a way that the mercury could flow from the former into the small separatory. By opening the stop-cocks of the two separataries partially and the other two cocks completely the gas was readily transferred. The rubber bulb, being in direct communication with the two vessels, indicated, by its dilations or contractions, the adjustments necessary to secure an equal flow of mercury through the two stop-cocks. It served, at the same time, to keep the pressure equal to the barometric pressure. Constant temperature during transference was secured by placing the bottle in a vessel of water kept at  $25^{\circ}$  ( $\pm 1^{\circ}$ ).

The main series of experiments was made with solutions of bicarbonate and carbonate in various proportions and of decinormal strength with respect to sodium. Such solutions, of approximately the desired composition, were made from separate solutions of bicarbonate and carbonate, each having the same concentration of sodium.

The method of calculating the equilibrium constant from the analytical data can doubtless best be shown by an example (Experiment 1, Table II):

## ANALYSIS OF THE GAS

Volume of the separatory	-	-	-	-	550 c.c.
Barometric pressure	-	-	-	-	743 mm.
Excess pressure in equilibrium bottle	-	-	-	-	2 mm.
Total pressure in equilibrium bottle	-	-	-	-	745 mm.
Temperature of the gas in the separatory immediately before analysis	-	-	-	-	28.2°

Added 10 c.c. dilute barium hydroxide (= 2.15 c.c. decinormal hydrochloric acid). 10 c.c. of gas escaped. After absorption of the carbon dioxide, the excess of alkali required 1.44 c.c. of acid,<sup>23</sup> phenolphthalein being used as indicator.

$$2.15 - 1.44 = 0.71 \text{ c.c. decinormal acid.}$$

$$1 \text{ c.c. of decinormal acid} = 1.12 \text{ c.c. of carbon dioxide gas at } 0^\circ \text{ and } 760 \text{ mm.}$$

$$0.71 \times 1.12 = 0.79 \text{ c.c.} = \text{the volume of the carbon dioxide at } 0^\circ \text{ and } 760 \text{ mm.}$$

$$\frac{0.79 \times (273 + 28) \times 760}{540 \times 273 \times 743} = \text{the fraction of the total pressure exerted by the carbon dioxide.}$$

The total pressure of the vapor phase is  $\frac{745}{760}$  of one atmosphere.

Therefore the partial pressure of the carbon dioxide,  $P$ , is

$$\frac{0.79 \times 301 \times 745}{540 \times 273 \times 743} = 0.00161 \text{ atmosphere.}$$

From which we get:

$$kP = 0.0338 \times 0.00161 = C_4.$$

#### ANALYSIS OF THE CARBONATE SOLUTION

The exact concentration of the carbonate solution was found from the titration in presence of methyl orange:

9.94 c.c. of carbonate solution required

9.93 c.c. of decinormal acid.

Therefore the solution was .0999 normal with respect to sodium.

To determine the percentage of bicarbonate 9.94 c.c. of the carbonate solution was run into 7.49 c.c. of decinormal barium hydroxide and 2 c.c. of approximately seminormal barium chloride. The excess of alkali required 0.70 c.c. of decinormal hydrochloric acid. The difference, 6.79 c.c., was therefore the amount of barium hydroxide required to neutralize all of the carbon dioxide in the solution in excess of that needed to form the normal carbonate with the sodium. The larger part of this was united to form bicarbonate, but a small amount was present as free carbonic acid. The concentration of the free carbonic acid, considered as monobasic, has just been found to be  $0.00161 \times 0.0338$  normal. In neutralizing barium hydroxide it is dibasic. The volume of decinormal barium hydroxide required to neutralize 10 c.c. of carbonic acid of this strength is therefore

$$10 \times 10 \times 2 \times 0.00161 \times 0.0338 \text{ c.c.} = 0.01 \text{ c.c.,}$$

the correction for the free carbonic acid in the solution. Therefore 6.78 c.c. is the amount of barium hydroxide required for the bicarbonate alone; and  $\frac{6.78}{9.93}$ , or 68.3 per

<sup>23</sup> In titrations requiring small volumes of solution (as in this case), a special gravimetric burette, holding about 4 cc., was employed, and the amount of solution used deter-

mined by weighing. Much greater accuracy was thereby obtained.

cent. of the sodium, was in the form of bicarbonate. The duplicate analysis gave 68.1 per cent. Mean 68.2. Therefore

$$x = 0.682 . \quad 1 - x = 0.318 . \quad P = 0.00161 . \quad k = 0.0338 .$$

Substituting these values in the formula

$$\frac{2x^2C}{kP(1-x)} = K$$

we get

$$\frac{2 \times 0.682^2 \times 0.0999}{0.0338 \times 0.00161 \times 0.318} = 5320 = K .$$

The following table shows the results obtained with solutions decinormal with respect to sodium. The fourth column contains the apparent percentages,  $A$ , of bicarbonate. These values, corrected in the manner described for the free carbonic acid in the solution, give the real percentages of bicarbonate as contained in the fifth column:

TABLE II

	Baro- metric Pressure	Excess Pressure	$A$	100 $x$	100 $P$	$K \times 10^{-2}$
1.....	743	2	68.3	68.2	.161	53.2
2.....	740	2	69.1	69.0	.159	57.0
3.....	745	2	76.2	76.0	.259	54.6
4.....	742	3	77.7	77.5	.294	53.6
5.....	743	10	78.3	78.1	.322	51.0
6.....	743	3	82.1	81.8	.404	53.6
7.....	745	0	96.6	96.1	2.23	49.0
8.....	740	24	103.6	98.5	7.49	51.0
						Mean 52.9

The above results represent all experiments, carried to completion, on solutions of this concentration, after the method, as described, had been fully worked out. An inspection of the formula for the equilibrium constant will show that for certain proportions of the reacting substances very small errors in the analysis will cause large errors in the value of  $K$ . This is obviously the case, on the one hand, when the percentage of bicarbonate becomes so high that  $1-x$  approaches zero, and, on the other hand, when  $P$  becomes very small on account of the large proportion of carbonate. The most favorable condition for the accurate determination of  $K$  is found when  $x$  lies between 75 and 85 per cent. In accordance with this it may be seen, from the table, that the smallest variations in the value of  $K$  are found between these limits. The average value of  $K$  as given by these four determinations is  $53.2 \times 10^2$ . This is the most probable value of the constant.

Walker<sup>24</sup> has calculated, from the ionization constant of carbonic acid, the theo-

<sup>24</sup> *Zeitschr. f. physik. Chemie*, Vol. XXXII (1900), p. 137.



retical value of the hydrolytic dissociation of sodium bicarbonate. This was done by means of the well-known formula of Arrhenius,<sup>28</sup> which here becomes

$$K_1 = \frac{K_{H_2O}}{K_{H_2CO_3}} .$$

The closely agreeing results of the determinations of the magnitude of the dissociation constant of water by Ostwald,<sup>28</sup> Wys,<sup>29</sup> Arrhenius,<sup>28</sup> and Kohlrausch and Heydweiller<sup>30</sup> give as an average value<sup>30</sup>

$$K_{H_2O} = 1.2 \times 10^{-14} .$$

The constant of carbonic acid, as measured by the conductivity method by Walker and Cormack,<sup>31</sup> gave the value,

$$K_{H_2CO_3} = 3.04 \times 10^{-7} ,$$

the concentration of the carbonic acid being expressed in gram-molecules per liter. Therefore

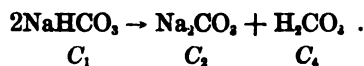
$$K_1 = \frac{1.2 \times 10^{-14}}{3.04 \times 10^{-7}} = 3.95 \times 10^{-8} .$$

But

$$K_1 = \frac{C_{NaOH} \times C_{H_2CO_3}}{C_{NaHCO_3}} . \quad (\text{See equation 1.})$$

From these relationships Walker obtained for the hydrolytic dissociation of a decinormal solution of sodium bicarbonate 0.06 per cent. This value is really fictitious, for the reason shown on p. 4. This is clearly seen when we consider that for 0.06 per cent. hydrolysis the concentration of the carbonic acid in the decinormal bicarbonate solution would be 0.00006 normal molecular. The concentration of carbonic acid in water saturated at atmospheric pressure at 25° is 0.0338 normal molecular. Therefore the pressure of 0.00006 normal solution is only 0.0018 atmosphere. From Table II it is seen that a decinormal solution containing but 96 per cent. of bicarbonate gives a partial pressure of more than twelve times this amount.

The actual concentrations of the free carbonic acid and of the sodium hydroxide in a decinormal solution made from solid bicarbonate and water in such a way that no carbon dioxide is allowed to escape may readily be calculated. Since the amount of free hydroxide in such a solution is extremely minute, as is shown below, we may represent the net result of the hydrolytic dissociations thus:



If

$$\frac{x}{V} = C_1 , \quad \frac{1-x}{2V} = C_2 = C_4 ,$$

$$K = \frac{C_1^2}{C_2 C_4} = \frac{4x^2}{(1-x)^2} = 5290 .$$

<sup>28</sup> *Zeitschr. f. physik. Chemie*, Vol. V (1890), p. 17.

<sup>29</sup> *Ibid.*, Vol. XI (1893), p. 521.

<sup>30</sup> *Ibid.*, p. 492.

<sup>31</sup> *Ibid.*, p. 827.

<sup>32</sup> *Ibid.*, Vol. XIV (1894), p. 317.

<sup>33</sup> VAN'T HOFF, *Vorlesungen*, Vol. I, p. 127.

<sup>34</sup> *Jour. Chem. Soc. (London)*, Vol. LXXVII (1900), p. 5.

From which

$$x = 0.9732, \text{ and } 1 - x = 0.0268.$$

In such a solution 2.68 per cent. of the bicarbonate would have been converted into carbonate and free carbonic acid. It will be observed, further, that the equation involving  $x$  does not contain  $V$ . The value of  $x$  is therefore independent of the dilution. For a decinormal solution

$$C_4 = \frac{1-x}{2 \times 10} = 0.00134.$$

$$P = \frac{0.00134}{0.0338} = 0.0396.$$

From which it follows that a decinormal solution of sodium bicarbonate will evolve carbon dioxide if the partial pressure of the gas is less than 4 per cent. of one atmosphere. Now for the bicarbonate solution we have

$$K_1 = \frac{C_{\text{NaOH}} \times C_{\text{H}_2\text{CO}_3}}{C_{\text{NaHCO}_3}} = 3.95 \times 10^{-8},$$

$$C_{\text{H}_2\text{CO}_3} = C_4 = 0.00134,$$

$$C_{\text{NaHCO}_3} = C_1 = 0.097.$$

Therefore

$$C_{\text{NaOH}} = 2.9 \times 10^{-8}.$$

This value, which is but one-twentieth of that calculated by Walker,<sup>22</sup> represents the true degree of alkalinity of a freshly prepared decinormal solution of sodium bicarbonate. Solutions consisting of mixtures of bicarbonate and carbonate contain, of course, more free sodium hydroxide. Thus the concentration of sodium hydroxide in solution 1, Table II, is readily calculated to be  $5.0 \times 10^{-5}$ .

According to equation 3, p. 5, we have

$$K_2 = KK_1,$$

where  $K_1$  is the constant for the hydrolytic dissociation of sodium bicarbonate and  $K_2$  that for the carbonate.

$$K = 5290. \quad K_1 = 3.95 \times 10^{-8}.$$

From which

$$K_2 = 2.08 \times 10^{-4}.$$

The percentage of hydrolytic dissociation of sodium carbonate may be calculated from the value of this constant by means of the formula

$$K_2 = \frac{m^2}{(1-m)V} = 2.08 \times 10^{-4},$$

where  $m$  is the fraction of the carbonate existing as hydroxide and bicarbonate. For a solution containing one-half gram-molecule in ten liters,  $m = 0.062$ . That is, in a solution decinormal with respect to sodium, 6.2 per cent. of the sodium carbonate is

<sup>22</sup> *Zeitschr. f. physik. Chemie*, Vol. XXXII (1900), p. 137.

split by the action of water into sodium hydroxide and sodium bicarbonate. The concentration of the hydroxide in such a solution is 0.0031 normal.

The hydrolytic dissociation of sodium carbonate has been studied by Shields.<sup>23</sup> The amount of free base was calculated from the velocity of saponification of ethyl acetate at 24.2°. Shields gives as the value of the constant for a 0.0477 gram-molecular solution<sup>24</sup>

$$(100) K_1 = 0.02382 ,$$

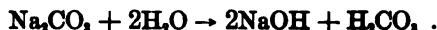
that is,

$$K_1 = 2.38 \times 10^{-4} ,$$

while I find, at 25°,

$$K_1 = 2.08 \times 10^{-4} .$$

The agreement *seems* to be very good. But it appears, upon critical inspection of the method of calculation, that Shields has assumed that each molecule of carbonate hydrolysed gives two molecules of hydroxide, thus:



As this is not the case to any appreciable extent, it would seem that the method of calculation is not correct. The results of these experiments of Shields have been recalculated for the 0.0477 molecular solution on the basis of a dissociation into one molecule each of hydroxide and bicarbonate. The average value of  $K_1$ , thus found is  $1.04 \times 10^{-4}$ , from which the percentage of carbonate molecules changed to hydroxide and bicarbonate is calculated to be 4.6. My own experiments lead to the value 6.2 per cent. for a 0.05 gram-molecular solution. Shields calculates that 4.87 per cent. of the carbonate molecules each give two of hydroxide.

Walker and Cormack<sup>25</sup> have shown that the ionization of carbonic acid is chiefly into H and  $\text{HCO}_3$ . The constant for the second H ion is fixed by the relation

$$C_{\text{H}} \times C_{\text{CO}_3} = K_2 C_{\text{HCO}_3} .$$

The value of  $K_2$  has been calculated by Bodländer<sup>26</sup> from the work of Shields. Bodländer finds

$$K_2 = 1.295 \times 10^{-11} .$$

$K_2$  may be calculated from any experiment recorded in Table II. Experiment 6 may be chosen, as it probably represents the most favorable composition for the accurate analytical determination of all the factors,  $x$ ,  $1 - x$ , and  $P$ :

$$x = 0.818 . \quad 1 - x = 0.182 . \quad P = 0.00404 .$$

Since practically all of the  $\text{HCO}_3$  ions come from the sodium bicarbonate,

$$C_{\text{HCO}_3} = 0.1 \times 0.818 \times \alpha ,$$

where  $\alpha$  is the degree of ionization of the bicarbonate. By an interpolation of

<sup>23</sup> *Zeitschr. f. physik. Chemie*, Vol. XII (1898), p. 167.

<sup>25</sup> *Jour. Chem. Soc. (London)*, Vol. LXXVII (1900), p. 5.

<sup>24</sup> Shields calls this  $K$ , and not  $100K$ , as I have written it and as it really should be. See Van Laar, *ibid.*, Vol. XII (1898), p. 745.

<sup>26</sup> *Loc. cit.*

Walker and Cormack's figures the degree of ionization of 0.0818 molecular sodium bicarbonate is calculated to be 0.81. Therefore

$$C_{\text{HCO}_3} = 0.0818 \times 0.81 = 6.63 \times 10^{-3}.$$

Similarly the concentration of  $\text{CO}_3$  depends upon the concentration of the sodium carbonate and on its degree of ionization. If we consider the ionization of the carbonate to be the same as that of sodium sulphate,<sup>27</sup> at equal concentration, we have for a  $\frac{1}{2} \times 0.0182$  molecular solution 0.83. Therefore

$$C_{\text{CO}_3} = \frac{1}{2} \times 0.0182 \times 0.83 = 7.55 \times 10^{-4}.$$

Now

$$C_{\text{H}} \times C_{\text{HCO}_3} = 3.04 \times 10^{-7} \times C_{\text{H}_2\text{CO}_3}.$$

$$C_{\text{H}_2\text{CO}_3} = 0.00404 \times 0.0338 = 1.37 \times 10^{-4}.$$

$$C_{\text{H}} = \frac{3.04 \times 1.37 \times 10^{-11}}{6.63 \times 10^{-3}} = 6.28 \times 10^{-10}.$$

Therefore

$$K_3 = \frac{6.28 \times 7.55 \times 10^{-13}}{6.63 \times 10^{-3}} = 7.16 \times 10^{-11}.$$

The considerable difference between this result and that of Bodländer is in a large measure due to Shields's error in calculating the hydrolysis of sodium carbonate. A comparison of  $K_3$  with the ionization constant for the first H ion of carbonic acid shows that the latter is 4,250 times as large as the former. Since the degrees of ionization, at the same concentration, of two weak acids are proportional to the square roots of their ionization constants, the degree of ionization of the first H ion of carbonic acid is  $\sqrt{4250}$ , or 65 times as great as that of the second H ion. The ionization constants of several very weak acids are all greater than  $K_3$ , as the following figures of Walker and Cormack show:

	Const. $\times 10^{-10}$
Carbonic acid (first H ion) - - - - -	3040
Hydrogen sulphide - - - - -	570
Boric acid - - - - -	17
Prussic acid - - - - -	13
Phenol - - - - -	1.3

I find: Carbonic acid (second H ion) =  $K_3$  - - - 0.72

In very dilute solutions sodium bicarbonate is highly ionized into Na and  $\text{HCO}_3$ . The ionization constant  $K_4$  of sodium bicarbonate, as an acid, is therefore practically equal to  $K_3$  for such solutions. Sodium bicarbonate may also split into H and  $\text{NaCO}_3$  ions; and it is probable that the constant of this ionization is about the same as  $K_3$ . If this is the case,  $K_4$  in solutions not dilute is also nearly equal to  $K_3$ . The constant  $K_4$  may also be obtained, at least approximately, by dividing the constant for the dissociation of water by  $K_3$ , the constant for the hydrolytic dissociation of sodium carbonate.

<sup>27</sup> KOHLRAUSCH, *Wiedemann's Annalen*, Vol. XXVI (1885), p. 161.

$$K_1 = \frac{K_{H_2O}}{K_2} = \frac{1.2 \times 10^{-4}}{2.08 \times 10^{-4}} = 0.58 \times 10^{-10},$$

a quantity which does not differ greatly from the value of  $K_2$  found above.

The formula for the equilibrium of the system composed of carbon dioxide and the aqueous solution of the carbonates of sodium has been derived upon the supposition that the solution was dilute. It is not to be expected, therefore, that the value of  $K$  would be the same for concentrated solutions. Deviations from the simple relationships, which hold for dilute solutions, may be due to three causes: (1) the partial failure of the law of mass action; (2) the change in the ratio of the degrees of ionization, as salts, of the two carbonates; (3) a smaller solubility of carbon dioxide in the stronger solutions than in the dilute solutions, where the solubility has been assumed to be the same as in pure water. Equilibrium experiments with 0.3 normal (Table III) and with normal solutions (Table IV) actually gave lower values for  $K$  than that found for decinormal solutions; but there is good agreement between various determinations at the same concentrations. The following tables show the results obtained:

TABLE III

	Barometric Pressure	Excess Pressure	$A$	$100x$	$100P$	$K \times 10^{-10}$
1.....	747	0	58.0	57.9	.319	44.6
2.....	748	0	68.0	67.9	.583	43.8
3.....	745	0	77.1	78.9	1.044	43.6
4.....	740	3	85.6	85.2	1.07	44.4
5.....	740	3	89.4	88.8	2.76	45.4
6.....	743	2	93.6	92.6	4.51	45.8
						Mean 44.6

TABLE IV

	Barometric Pressure	Excess Pressure	$A$	$100x$	$100P$	$K \times 10^{-10}$
1.....	744	0	76.1	75.8	4.36	32.4
2.....	737	6	81.4	81.0	6.24	32.8
3.....	742	2	86.7	86.0	10.21	30.6
4.....	743	1	91.3	90.2	16.82	29.2
						Mean 31.2

#### EQUILIBRIUM IN CARBONATE-BICARBONATE SOLUTIONS IN CONTACT WITH NATURAL AIR

The determination of the amount of carbon dioxide in the air has been the subject of a very large number of extended researches. In a recent paper<sup>22</sup> Letts and Blake have given the results of their own very careful experiments, made at Belfast, together with a tabulated summary of several thousand analyses made by many observers in

<sup>22</sup> *Scientific Proc. Royal Dublin Soc.*, 1900, pp. 107-279.

various parts of the world. This summary clearly shows that the percentage of carbon dioxide in air (collected out of doors) is by no means constant. But, if analyses of air from the central portions of crowded cities be excluded, all except a small portion of the results fall between the limits of 2.5 and 3.5 parts by volume of carbon dioxide in 10,000 of dry air. The average value is probably not far from three parts in 10,000.

Taking the amount of carbon dioxide as fixed at this value, we may calculate the theoretical composition of carbonate-bicarbonate solutions when in equilibrium with ordinary air, at a temperature of 25°. In the formula

$$\frac{2x^2C}{kP(1-x)} = K,$$

$P$ , the partial pressure of the carbon dioxide, is equal to

$$\frac{0.0003 H}{760},$$

where  $H$  is the barometric pressure (corrected for aqueous tension) at the time of the experiment. Transposing we have

$$\frac{x^2}{1-x} = \frac{K \times H \times 0.0338 \times 0.0003}{2 \times 760 \times C}. \quad (4)$$

In the work of Cameron and Briggs, which has been previously mentioned, one series of experiments on the carbonates of sodium was carried out at 25°. The theoretical results, as indicated by the above formula, have been calculated and compared with the values obtained by these chemists. In making the calculation  $H$ , the corrected barometric pressure, has been assumed equal to 730 mm. The value of  $K$  for dilute solutions is practically  $53 \times 10^2$ . The values of  $K$  used in calculating the theoretical values of  $x$ , for the more concentrated solutions, were obtained by interpolation from the average values of  $K$  as found in Tables III and IV.

The comparison of Cameron and Briggs's results with the theoretical values is shown in the following table:

TABLE V

Concentration of Sodium in Gram-Atoms per Liter	Value of $K \times 10^{-2}$ Used	PERCENTAGE, $x$ , OF SODIUM EXISTING AS BICARBONATE	
		Found by C. and B.	Calculated
0.0044	53	91.3	87
0.0143	53	80.0	72
0.0562	53	62.7	49
0.2248	48	40.7	28
0.8847	33	36.0	13

The differences between the calculated results and those actually found by Cameron and Briggs were so great, especially for the more concentrated solutions, that I was led to make a few experiments of the same nature.

The results of a preliminary experiment may serve to show in a roughly quantitative way the rate of change of composition of an approximately decinormal solution of sodium bicarbonate when a stream of air (taken from out of doors) is drawn through it at the ordinary temperature of the laboratory:

TABLE VI

Time in Hours	Percentage of Sodium Existing as Bicarbonate
0	95
0.25	88
0.87	80
1.17	76
2.00	75
28	45
32	44
48	43
72	40
144	40

A state of equilibrium is therefore reached after about three days.

In a second similar experiment air was drawn through a solution of sodium carbonate having the same concentration of sodium as the preceding bicarbonate solution. The result was as follows:

TABLE VII

Time in Hours	Percentage of Sodium Existing as Bicarbonate
0	0.3
5.5	5.0
17	17
24	26
70	40

The same state of equilibrium is reached, of course, whether one starts with a carbonate or a bicarbonate solution.

In the following experiments about 200 cc. of the bicarbonate solution was placed in a large Drechsel wash bottle, which was immersed in a thermostat kept at 25° ( $\pm 0.01^\circ$ ). The bottle containing the bicarbonate solution was connected on the one side with an aspirating pump and on the other with a wash bottle containing pure water. Pure air (from out of doors) was supplied by a tube connected with the wash bottle.

The pressure in the bottle containing the bicarbonate solution was measured by means of a manometer. It was always 8 to 10 mm. (of mercury) less than the atmospheric pressure. The corrected pressure, as given in Tables VIII, IX, and X, is the barometric pressure, corrected for aqueous tension and for the diminution just mentioned.

The following results were obtained in this way, starting with a decinormal sodium bicarbonate solution:

TABLE VIII

Time in Days	Concentration of Sodium in Gram- Atoms per Liter	Corrected Pressure	PERCENTAGE OF SODIUM EXISTING AS BICARBONATE	
			Found	Calculated
0	0.100	...	95	$(K = 53 \times 10^3)$
2	.....	...	44.1	
3	0.101	710	40.6	
3.2	0.101	713	40.8	

The figures in the last column were obtained by means of equation 4, p. 17,  $K$  being taken as  $53 \times 10^3$ . The fact that the solution slowly became slightly more concentrated shows that the air entering the solution was not completely saturated with water vapor.

The results with more concentrated solutions are given in Tables IX and X:

TABLE IX

Time in Days	Concentration of Sodium in Gram- Atoms per Liter	Corrected Pressure	PERCENTAGE OF SODIUM EXISTING AS BICARBONATE	
			Found	Calculated
(FIRST SERIES)				
0	0.300	...	33.0	(K = 45×10 <sup>3</sup> )
1	. . . .	...	26.2	
2	.....	...	24.8	
3	0.312	710	23.9	
(SECOND SERIES)				
0	0.300	...	22.7	.....
1	0.306	711	23.1	23.2
2	0.314	714	22.5	22.9

TABLE X

Time in Days	Concentration of Sodium in Gram- Atoms per Liter	Corrected Pressure	PERCENTAGE OF SODIUM EXISTING AS BICARBONATE	
			Found	Calculated
0	1.00	...	13.8	$(K = 31 \times 10^3)$
2	1.03	710	12.2	
3	0.98 <sup>20</sup>	714	12.9	
5	0.98	715	12.9	

<sup>20</sup> The solution was slightly diluted at the end of the second day.



In the second series of experiments, Table IX, the solution at the start contained a smaller amount of bicarbonate than at the end of one day. This indicates that an absorption of carbon dioxide from the air must have occurred.

A variation in the amount of carbon dioxide in the air necessarily causes a variation in the percentage of bicarbonate in a solution of given concentration. Calculation shows that an increase of 10 per cent. of the total carbon dioxide in the air would cause an increase, in a decinormal solution, of only 1.4 per cent. of the sodium existing as bicarbonate. In Table VIII, taking the amount of carbon dioxide in the air as 3 parts in 10,000, the percentage of bicarbonate was calculated to be 39.0. For 3.3 parts of carbon dioxide in 10,000 of air the amount would therefore be 40.4 per cent. The increase in bicarbonate for a 10 per cent. increase of carbon dioxide for a 0.3 normal solution would be 0.9 per cent., and for a normal solution but 0.5 per cent.

The ordinary variations in carbon-dioxide content of air rarely exceeds  $\pm 10$  per cent. from the assumed value, 3 parts in 10,000 (except in certain localities). Such variations would consequently affect but slightly the state of equilibrium of solutions in contact with the air. In accordance with this it is seen that the agreement between the observed values and those calculated upon the basis of the average concentration of carbon dioxide in the air is very good; and, further, that the small existing differences are no greater than may be accounted for by changes in the amount of atmospheric carbon dioxide.<sup>40</sup>

These experiments are therefore in complete accord with those of the main series. This being the case, it is difficult to understand the cause of the widely different results obtained by Cameron and Briggs. A possible source of error may be suggested. The authors do not state whether the air used was drawn from out of doors or was simply indoor air. In the latter case the amount of carbon dioxide may have been very much higher than in normal air. This would lead to a displacement of the equilibrium in the direction found.

#### SUMMARY

1. It has been shown experimentally that the equilibrium in the system composed of sodium carbonate, sodium bicarbonate, carbon dioxide, and water is governed, at constant temperature, by the formula,

$$\frac{2x^2 C}{kP(1-x)} = K ,$$

in which the symbols have the following significance:  $x$  is the fraction of sodium in the form of bicarbonate;  $1-x$  is the fraction in the form of carbonate;  $C$  is the concentration of the sodium in gram-atoms per liter;  $k$  is the solubility coefficient of

<sup>40</sup>As has been mentioned, analyses seem to show that the percentage of carbon dioxide in the air of the crowded portions of large cities is considerably higher than normal. The laboratory of the University of Chicago, where the ex-

periments here described were made, is six or seven miles distant from the central crowded parts of the city, and but a mile from Lake Michigan. In consequence the air used was probably of about normal composition.

carbon dioxide in water at the equilibrium temperature; and  $P$  is the partial pressure of the carbon dioxide. For dilute solutions  $K = 53 \times 10^3$ .

2. The degree of hydrolytic dissociation of sodium bicarbonate, as calculated from the ionization constant of carbonic acid, is rendered fictitious by the fact that the bicarbonate is itself a weak acid. The acid carbonate reacts with the free hydroxide produced by the primary hydrolysis, and, on account of its relatively great concentration as compared with that of the free carbonic acid, greatly displaces the equilibrium that would be expected to result from the first action of water on the bicarbonate. For a decinormal solution the free hydroxide is thereby reduced to about one-twentieth of the amount calculated by Walker.

3. The degree of hydrolytic dissociation of sodium carbonate was calculated from the equilibrium experiments. In a solution containing one-half gram-molecule of sodium carbonate in ten liters, 6.2 per cent. of the carbonate exists as hydroxide and bicarbonate.

4. The ionization constant of sodium bicarbonate, as an acid, is nearly equal to the constant of the second H ion of carbonic acid. The latter constant is equal to  $7.2 \times 10^{-11}$ , which is but  $\frac{1}{4500}$  as great as that of the first H ion of carbonic acid.

5. It has long been known that bicarbonate solutions lose carbon dioxide in contact with the air, giving a mixture of carbonate and bicarbonate. The state of equilibrium reached depends on the concentration of the solution, and may be calculated by means of the equilibrium formula by considering the amount of carbon dioxide in the air as known.

6. The formula give above will hold equally well for solutions of the carbonates of any highly ionized monacid bases. It is probable that it may also be applied to salts of such acids as hydrogen sulphide and sulphurous acid.



# **INVARIANTS OF DIFFERENTIAL QUANTICS**



# INVARIANTS AND COVARIANTS OF QUADRATIC DIFFERENTIAL QUANTICS OF $n$ VARIABLES

H. MASCHKE

IN the article<sup>1</sup> "A New Method of Determining the Differential Parameters and Invariants of Quadratic Differential Quantics, I have shown that the application of a certain symbolic method leads very readily to the formation of expressions remaining invariant with respect to the transformation of quadratic differential quantics. While, however, in that article I have confined the work essentially to the case of two independent variables, in the following an attempt will be made to treat the general case of  $n$  independent variables.

## §1. DEFINITIONS. THE FUNDAMENTAL THEOREM

To the given quadratic differential quantic

$$A = \sum_{i,k=1}^n a_{ik} dx_i dx_k, \quad (1)$$

with  $x_1, x_2, \dots, x_n$  as independent variables, and the  $a_{ik}$  ( $a_{ki}=a_{ik}$ ) as functions of these variables, we apply the transformation

$$x_i = x_i(y_1, y_2, \dots, y_n) \quad (i=1, 2, \dots, n), \quad (2)$$

and obtain

$$A = A' = \sum_{i,k=1}^n a'_{ik} dy_i dy_k. \quad (3)$$

Since the differentials  $dy$  are connected linearly with the differentials  $dx$  by means of the formulas

$$dx_i = \sum_{k=1}^n \frac{\partial x_i}{\partial y_k} dy_k \quad (i=1, 2, \dots, n), \quad (4)$$

we have at once, denoting the discriminant of  $A$  by  $|a_{ik}|$ ,

$$|a'_{ik}| = r^2 |a_{ik}|, \quad (5)$$

where  $r$  denotes the determinant of the linear substitution, viz.:

$$r = \left| \frac{\partial x_i}{\partial y_k} \right|. \quad (6)$$

Let now  $u, v, \dots$  be any (arbitrary) functions of  $x_1, x_2, \dots, x_n$ ;  $u', v', \dots$  the same functions after application of the transformation (2), then we call every function of the coefficients  $a_{ik}$  and their derivatives, and of  $u, v, \dots$  and their derivatives, an

<sup>1</sup> *Transactions of the American Mathematical Society*, Vol. I, No. 2, pp. 197-204.

*invariant expression* of the quadratic differential quantic  $A$ , if the expression remains the same, whether formed with the old quantities  $a_{ik}, u, v, \dots$  and their derivatives with respect to  $x$ , or with the new quantities  $a'_{ik}, u', v', \dots$  and their derivatives with respect to  $y$ .

From  $u' = u$  it follows, for instance, that every arbitrary function of  $x$  is an invariant expression of  $A$ .

If such an invariant expression involves one or more arbitrary functions  $u, v, \dots$  and their derivatives, it is called a *differential parameter*; if it involves no such functions, if it is therefore a function of the  $a_{ik}$  and their derivatives alone, it is called an *invariant proper*.<sup>3</sup>

By the *order* of an invariant expression we shall understand the order of the highest derivative appearing in it.

Suppose now  $F^{(1)}, F^{(2)}, \dots, F^{(n)}$  are any  $n$  invariant expressions of  $A$ , then we have

$$F'^{(i)} = F^{(i)} \quad (i=1, 2, \dots, n),$$

and also

$$\sum_{k=1}^n \frac{\partial F'^{(i)}}{\partial y_k} dy_k = \sum_{k=1}^n \frac{\partial F^{(i)}}{\partial x_k} dx_k \quad (i=1, 2, \dots, n).$$

It now follows at once

$$\left| \frac{\partial F'^{(i)}}{\partial y_k} \right| = r \cdot \left| \frac{\partial F^{(i)}}{\partial x_k} \right|,$$

and therefore with reference to (5)

$$|a'_{ik}|^{-\frac{1}{2}} \cdot \left| \frac{\partial F'^{(i)}}{\partial y_k} \right| = |a_{ik}|^{-\frac{1}{2}} \cdot \left| \frac{\partial F^{(i)}}{\partial x_k} \right|. \quad (7)$$

This equation defines the last expression as an invariant expression, and so we have the fundamental theorem:

"If  $F^{(1)}, F^{(2)}, \dots, F^{(n)}$  are any  $n$  invariant expressions of  $A$ , then

$$|a_{ik}|^{-\frac{1}{2}} \cdot \left| \frac{\partial F^{(i)}}{\partial x_k} \right|$$

is again an invariant expression of  $A$ ."

## §2. CHOICE OF CONVENIENT NOTATIONS

Since we shall have in the following to compute continuously with expressions of the type (7), a shorter notation is indispensable. Let us first agree to indicate differentiation by subscripts.  $F$  being any quantity whatever, we write

$$\frac{\partial F}{\partial x_i} = F_i. \quad (8)$$

<sup>3</sup> Cf. LUIGI BIANCHI, *Vorlesungen über Differentialgeometrie*; autorisierte deutsche Uebersetzung von MAX LUKAT (Leipzig, 1899), p. 39.

We further denote the reciprocal value of the (positive) square root of the discriminant of  $A$ —which we always suppose to be different from zero—by the single letter  $\beta$ :<sup>3</sup>

$$\beta = |a_{ik}|^{-\frac{1}{2}}. \quad (9)$$

The functional determinant of any  $n$  quantities  $F^{(i)}$  ( $i=1, 2, \dots, n$ ) (for which we simply write  $F^i$ , if there is no interference with the usual notation of powers) will be denoted by

$$\{F^1, F^2, F^3, \dots, F^n\},$$

so that we have

$$\{F^1, F^2, F^3, \dots, F^n\} = \left| \frac{\partial F^{(i)}}{\partial x_k} \right|. \quad (10)$$

The product of  $\beta$  into such a functional determinant will be denoted by

$$(F^1, F^2, \dots, F^n) = \beta \{F^1, F^2, \dots, F^n\}. \quad (11)$$

The quantities we have almost exclusively to deal with in the sequel are not the functional determinants themselves, but their products into  $\beta$ , and for this reason we use for the latter quantities the simpler symbol  $( )$  instead of  $\{ \}$ .<sup>4</sup>

Even this notation is for most cases too cumbersome. We write then simply

$$(F^1, F^2, \dots, F^n) = (F). \quad (12)$$

If it should be necessary to indicate the first, or the first two, three, etc., quantities of such a parenthesis distinctly, we write them in their proper places and let the last letter run out. For instance

$$\begin{aligned} (a) &\text{ means } (a^1, a^2, \dots, a^n), \\ (b, a) &\text{ means } (b, a^2, a^3, \dots, a^n), \\ (b, c, a) &\text{ means } (b, c, a^3, a^4, \dots, a^n). \end{aligned} \quad (13)$$

I give the last example in full:

$$(b, c, a) = \left| \begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{array} \right|^{-\frac{1}{2}} \cdot \left| \begin{array}{cccc} \frac{\partial b}{\partial x_1} & \frac{\partial b}{\partial x_2} & \dots & \frac{\partial b}{\partial x_n} \\ \frac{\partial c}{\partial x_1} & \frac{\partial c}{\partial x_2} & \dots & \frac{\partial c}{\partial x_n} \\ \frac{\partial a^{(1)}}{\partial x_1} & \frac{\partial a^{(1)}}{\partial x_2} & \dots & \frac{\partial a^{(1)}}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial a^{(n)}}{\partial x_1} & \frac{\partial a^{(n)}}{\partial x_2} & \dots & \frac{\partial a^{(n)}}{\partial x_n} \end{array} \right|.$$

It is understood that the letter  $a$  occurring in the coefficients  $a_{ik}$  of  $A$  has no connection whatever with the letter  $a$  occurring, for instance, in  $(b, c, a)$ .

<sup>3</sup> as in my previous paper.

<sup>4</sup> I therefore withdraw the suggestion made in my

paper, *loc. cit.*, p. 199, footnote, where the two parentheses  $( )$  and  $\{ \}$  were used in the reversed sense.



The quantities in a parenthesis ( ) should be separated by commas. If, however, no misrepresentation can occur, the commas may be omitted:

$$(bca) = (b, c, a), \text{ etc.}$$

The only notation left for the ordinary parenthesis in mathematical formulas is then [ ], since ( ) and { } have now both a specific meaning. Where there is no misunderstanding, also ( ) and { } may be used for ordinary parentheses.

In our new notation the last theorem of § 1 is now this: "If  $F^1, F^2, \dots, F^n$  are invariant expressions of  $A$ , then  $(F)$  is also an invariant expression of  $A$ ;" we shall call it an *invariantive constituent*.

### § 3. THE SYMBOLIC METHOD

If we define

$$f_i f_k = a_{ik}, \quad (14)$$

we have

$$A = \sum_{ik} a_{ik} dx_i dx_k = \left[ \sum_i f_i dx_i \right]^2. \quad (15)$$

The expression  $\sum_i f_i dx_i$  appears, if we think of the notation  $f_i = \frac{\partial f}{\partial x_i}$  agreed upon in (8) as the complete differential of a (symbolic) function of the  $n$  variables  $x$ .

If expressions of higher than the first dimensions in the coefficients  $a_{ik}$  are to be formed symbolically, we have to use different symbols  $f, \phi, \dots$ .

For instance

$$a_{12}^2 = f_1 f_2 \phi_1 \phi_2, \quad a_{11} a_{22} = f_1 f_1 \phi_2 \phi_2, \text{ etc.}$$

The symbolic functions  $f, \phi, \dots$  (we shall simply call them symbols) appear now, as every arbitrary function  $u, v, \dots$  of the variables  $x$  does, as invariant expressions of  $A$ .

If now we form invariantive constituents containing the symbols  $f, \phi, \dots$  or  $f^1, f^2, \dots$  and any number of arbitrary functions  $u, v, \dots$ , then every product of these constituents will represent an expression which, according to the fundamental theorem of § 1, will represent an expression which is *formally* invariant. But from the same reasoning as in algebra it follows that these products will at once represent *actual* invariant expressions, as soon as the symbols  $f, \phi, \dots$  occur in such connections as to permit actual meaning, *e. g.*, in the connection  $f_i f_k, \phi_i \phi_k$ , etc. The connections of this type, however, are not the only ones. Some of the elements of the different invariantive constituents may be constituents themselves, as for instance in the examples

$$((fa), b) \cdot (fc) \text{ or } ((fa), (fb), c),$$

where  $f$  denotes a symbol. In such cases also the higher derivatives of the symbol  $f$  will occur, and it is then the question, whether or not these combinations have actual meaning (they have it in the above examples). Thus we have the following theorem:

*Every product of invariantive constituents, the elements of which are symbols or arbitrary functions or both, or again invariantive constituents of the same character, represents an invariant expression of  $A$ , provided that every symbol occurs precisely twice and in such a connection as to permit actual interpretation in terms of the  $a_{ik}$  and their derivatives.*

With regard to these symbolic invariant expressions the following two theorems are evidently true:

*The value of an invariant expression in symbolic form is not changed if two equivalent symbols are interchanged, and:*

*An invariant expression in symbolic form vanishes if by the interchange of two equivalent symbols its sign is changed.*

Covariants can now also be formed easily. In the first place, the complete differential of every invariant expression represents immediately a linear covariant of  $A$ . Let us denote for simplicity

$$F_1 dx_1 + F_2 dx_2 + \dots F_n dx_n = F_x. \quad (16)$$

A single symbol and also any invariantive constituent with the subscript  $x$  represents then, at least formally, a linear covariant—let us call it a *covariantive constituent*. We therefore obtain covariants of any degree by forming products of these covariantive and invariantive constituents. Thus we have the theorem:

*Every product of covariantive and invariantive constituents represents an actual covariant of  $A$ , provided that every symbol occurs precisely twice and in such a connection as to permit actual interpretation in terms of the coefficients  $a_{ik}$  and their derivatives. The degree of the covariant is determined by the number of the covariantive constituents as factors in the product.*

#### §4. SOME IMPORTANT INVARIANT EXPRESSIONS

The simplest possible invariant proper is  $(f)^2$ , with  $f^1, f^2, \dots f^n$  as equivalent symbols. But this invariant reduces to a constant. To show it we compute first the product

$$P = f_1^1 f_2^2 \dots f_n^n \{f^1, f^2, \dots f^n\}.$$

If we multiply the first row of the functional determinant by  $f_1^1$ , the second by  $f_2^2$ , etc., we find at once

$$P = |a_{ik}| = \beta^{-1}.$$

If, on the other hand, we permute in  $P$  the equivalent symbols  $f^i$  in all possible ways, and add the so obtained products, which are all equal, together, we have

$$n! P = \{f^1, f^2, \dots f^n\}^2.$$

Since now

$$(f)^2 = \beta^2 \{f^1, f^2, \dots f^n\}^2,$$

we have the important formula

$$(f)^2 = n!. \quad (17)$$

by  $A_{ik}$ , the minor of the element  $A_{ik}$  in the determinant  $|a_{ik}|$ , (18)

by  $F^{i,k}$  the minor of the element  $f_k^i$  in the determinant  $\{f^1, f^2, \dots, f^n\}$ . (19)

$$(uf)^2 = \beta^2 \sum_{i,k} F^{1,i} \cdot F^{1,k} u_i u_k.$$
$$F^{1, \ell} F^{1, k} = (n-1)! f_1 f_2 \cdots f_{i-1} f_{i+1} \cdots f_n \cdot \begin{vmatrix} f_1^2, f_2^2, \cdots, f_{k-1}^2, f_{k+1}^2, \cdots, f_n^2 \\ \vdots \\ f_1^n, f_2^n, \cdots, f_{k-1}^n, f_{k+1}^n, \cdots, f_n^n \end{vmatrix},$$
$$F^{1,i} F^{1,k} = (n-1)! A_{ik},$$
$$F^{\lambda, i} F^{\lambda, k} = (n-1)! A_{ik} . \quad (20)$$
$$\frac{1}{(n-1)!} (uf)^2 = \beta^2 \sum_{i,k} A_{ik} \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_k}, \quad (21)$$
$$(uf)^2 = (u-1)! \Delta_1 u \quad . \quad (22)$$
$$\frac{1}{(n-1)!} (uf)(vf) = \beta^2 \sum_{ik} A_{ik} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_k},$$
$$(uf)(vf) = (u-1)! \Delta(u, v) \quad . \quad (23)$$

In all these examples of invariant expressions only the coefficients  $a_{ik}$  themselves occur, not their derivatives. In order to derive invariant expressions involving also derivatives of the  $a_{ik}$ , we have to express these in terms of our symbols.

$$\begin{aligned}\frac{\partial a_{ik}}{\partial x_i} &= f_i f_{ki} + f_k f_{ii}, \\ \frac{\partial a_{ki}}{\partial x_i} &= f_k f_{ii} + f_i f_{ki}, \\ \frac{\partial a_{ii}}{\partial x_i} &= f_i f_{ii} + f_i f_{ii},\end{aligned}\quad (24)$$

<sup>6</sup> BLANCHI, *loc. cit.*, p. 41.

which gives at once

$$f_i f_{kl} = \frac{1}{2} \left[ \frac{\partial a_{ik}}{\partial x_l} + \frac{\partial a_{il}}{\partial x_k} - \frac{\partial a_{kl}}{\partial x_i} \right]. \quad (25)$$

The expression on the right side is precisely Christoffel's<sup>1</sup> so-called triple index symbol denoted briefly by  $\begin{bmatrix} kl \\ i \end{bmatrix}$ .

We have thus the important theorem: *The symbolic product  $f_i f_{kl}$  has actual meaning for every system of values  $i, k, l$ ; it is equal to the triple index symbol*

$$f_i f_{kl} = \begin{bmatrix} kl \\ i \end{bmatrix}. \quad (26)$$

We see, then, further that the following combinations permit actual interpretation in the second derivatives of the  $a_{ik}$ :

$$\begin{aligned} f_{im} f_{kl} + f_i f_{klm} &= \begin{bmatrix} kl \\ i \end{bmatrix}_m, \\ f_{ir} f_{ks} - f_{kr} f_{is} &= \begin{bmatrix} ir \\ k \end{bmatrix}_s - \begin{bmatrix} is \\ k \end{bmatrix}_r, \\ f_r f_{iks} - f_s f_{ikr} &= \begin{bmatrix} ik \\ r \end{bmatrix}_s - \begin{bmatrix} ik \\ s \end{bmatrix}_r. \end{aligned} \quad (27)$$

The simplest symbolic invariant expression involving the second derivatives of the symbols  $f$  is  $((uf), f)$ . The computation which will be given at the end of §6 leads to the result

$$\frac{1}{(n-1)!} ((uf), f) = \beta^2 \cdot \sum_{ik} A_{ik} \frac{\partial^2 u}{\partial x_i \partial x_k} - \beta^4 \cdot \sum_{iklm} A_{ik} A_{lm} \begin{bmatrix} ik \\ l \end{bmatrix} \frac{\partial u}{\partial x_m}.$$

This expression is the second differential parameter  $\Delta_2 u$ .<sup>2</sup> Hence we have

$$((uf), f) = (n-1)! \Delta_2 u. \quad (28)$$

### §5. RELATIONS BETWEEN SYMBOLIC EXPRESSIONS

For our further computation with symbolic expressions it is necessary to deduce a number of fundamental relations. As a matter of convenience (not of definition) we shall as a rule denote symbols by the letters  $a, b, c, \dots, f, e, \dots$ , and any functions of the  $x$  (not necessarily symbols) by the letters  $u, v, \dots$ .

Differentiating formula (17) with regard to  $x_i$ , we have the fundamental formula

$$(f)(f)_i = 0. \quad (29)$$

Let now

$$[f] \text{ stand either for } (f) \text{ or for } (f)_i, \quad (30)$$

<sup>1</sup>CHRISTOFFEL, "Ueber die Transformation der homogenen Differentialausdrücke des zweiten Grades," *Crelle's Journal*, Vol. LXX, p. 48.

<sup>2</sup>BIANCHI, *loc. cit.*, p. 43.

<sup>3</sup>IDEM, *ibid.*, p. 47.

and let us form the symbolic product

$$f_1^1 (uf) [f] .$$

We find

$$f_1^1 \{uf\} [f] = \begin{vmatrix} f_1^1 u_1, u_2, \dots, u_n \\ f_1^1 f_1^1, f_2^2, \dots, f_n^n \\ \dots \\ f_1^1 f_1^n, f_2^n, \dots, f_n^n \end{vmatrix} \cdot [f] = \begin{vmatrix} f_1^1 u_1, u_2, \dots, u_n \\ 0, f_2^2, \dots, f_n^n \\ \dots \\ 0, f_2^n, \dots, f_n^n \end{vmatrix} \cdot [f]$$

because the product  $f_1^1 f_1^t [f]$  changes its sign if the equivalent symbols  $f^1$  and  $f^t$  are interchanged, and must therefore vanish; i. e.,

$$f_1^1 f_1^t [f] = 0 .$$

It now follows further

$$f_1^1 \{uf\} [f] = (n-1)! u_1 f_1^1 f_2^2 \dots f_n^n [f] = \frac{(n-1)!}{n!} u_1 \{f\} [f] ,$$

and

$$f_1^1 (uf) [f] = \frac{1}{n} u_1 (f) [f] .$$

The same method can be applied when we operate with  $f_i^1$  instead of  $f_1^1$ , so that we also have

$$f_i^1 (uf) [f] = \frac{1}{n} u_i (f) [f] . \quad (31)$$

If now we specify the symbol  $[f]$  according to (30) we obtain, changing the notation, the two equations

$$f_k (fa) (ua) = (n-1)! u_k , \quad (32)$$

$$f_k (fa)_i (ua) = 0 . \quad (33)$$

With  $v^1, v^2, \dots, v^n$  as arbitrary functions these formulas can at once be extended to

$$(fa) (fv) (ua) = (n-1)! (uv) , \quad (34)$$

$$(fa)_i (fv) (ua) = 0 . \quad (35)$$

A similar method serves to reduce the expression

$$f_1^1 f_2^2 (uvf) [f] .$$

We find

$$f_1^1 f_2^2 \{uvf\} [f] = \begin{vmatrix} f_1^1 u_1, f_2^2 u_2, u_3, \dots, u_n \\ f_1^1 v_1, f_2^2 v_2, v_3, \dots, v_n \\ f_1^1 f_1^1, f_2^2 f_2^2, f_3^3, \dots, f_n^n \\ \dots \\ f_1^1 f_1^n, f_2^2 f_2^n, f_3^n, \dots, f_n^n \end{vmatrix} \cdot [f] = \begin{vmatrix} f_1^1 u_1, f_2^2 u_2, u_3, \dots, u_n \\ f_1^1 v_1, f_2^2 v_2, v_3, \dots, v_n \\ 0, 0, f_3^3, \dots, f_n^n \\ \dots \\ 0, 0, f_3^n, \dots, f_n^n \end{vmatrix} \cdot [f] .$$

Further

$$f_1^1 f_2^2 \{uvf\} [f] = (u-2)! (u_1 v_2 - u_2 v_1) f_1^1 f_2^2 f_3^3 \dots f_n^n [f] = \frac{(u-2)!}{u!} (u_1 v_2 - u_2 v_1) \{f\} [f]$$

and

$$f_i^! f_k^! (uvf) [f] = \frac{1}{u(u-1)} (u_i v_k - u_k v_i) (f) [f] . \quad (36)$$

Specifying again the symbol  $[f]$ , and changing notation, we have finally

$$f_i \phi_k (f \phi a) (uva) = (n-2)! (u_i v_k - u_k v_i) , \quad (37)$$

$$f_i \phi_k (f \phi a)_i (uva) = 0 , \quad (38)$$

two equations, which can easily be extended to many others; I put down only the following three:

$$(f \phi w) (f \phi a) (uva) = 2(n-2)! (uvw) , \quad (39)$$

$$f_i (f \phi a)_k (\phi w) (uva) = 0 , \quad (40)$$

$$(f \phi w) (f \phi a)_k (uva) = 0 . \quad (41)$$

By differentiation we deduce from the above given formulas others which will be useful. Differentiating formula (31) with regard to  $x_k$ , and observing that

$$f_i^! (uf) [f]_k = \frac{1}{n} u_i (f) [f]_k$$

from the same reason which led to formula (31), we obtain

$$f_i^! (uf)_k [f] - \frac{1}{n} u_i (f)_k [f] = \frac{1}{n} u_{ik} (f) [f] - f_{ik}^! (uf) [f] ; \quad (42)$$

and finally by specifying  $[f]$  according to (30) and changing notation

$$f_i (ua)_k (fa) = (n-1)! u_{ik} - f_{ik} (ua) (fa) , \quad (43)$$

$$f_i (ua)_k (fa)_i = \frac{1}{n} (fa)_k (fa)_i u_i - f_{ik} (ua) (fa)_i . \quad (44)$$

The first of these equations (43) is of great importance. In every term of the last equation (44) connections of the form  $f_{ik} f_{rs}$  occur. Indeed, the different terms of this equation have no actual meaning. Nevertheless the formula is formally correct, and can be used for reduction work with safety. I omit to put down any of the great number of relations which can immediately be deduced from these two formulas.

#### §6. THE QUADRATIC COVARIANT EXPRESSION AND THE TRIPLE INDEX SYMBOLS OF THE SECOND KIND

From (43) we see that

$$f_i (ua)_k (fa) = f_k (ua)_i (fa) . \quad (45)$$

this quantity gives rise to the quadratic covariant expression

$$(fa) f_s (ua)_s = \sum_{ik} (fa) f_i (ua)_k dx_i dx_k , \quad (46)$$

whose coefficients can be shown to be equal to the "covariant second derivatives of the function  $u$ " in Bianchi's terminology.<sup>10</sup>

<sup>10</sup> Loc. cit., p. 46.

For this purpose we have first to compute the triple index symbol of the second kind<sup>11</sup>

$$\left\{ \begin{matrix} ik \\ l \end{matrix} \right\} = \beta^2 \sum_m A_{lm} \left[ \begin{matrix} ik \\ m \end{matrix} \right]. \quad (47)$$

From (20) and (26) we have

$$\begin{aligned} (n-1)! \left\{ \begin{matrix} ik \\ l \end{matrix} \right\} &= \beta^2 f'_{ik} F^{1,l} \cdot \sum_m f'_m F^{1,m} = \beta f'_{ik} F^{1,l} \cdot (f), \\ \left\{ \begin{matrix} ik \\ l \end{matrix} \right\} &= \frac{1}{(n-1)!} \beta f'_{ik} F^{1,l} \cdot (f), \end{aligned} \quad (48)$$

which is the required symbolic representation.

We derive from this equation

$$\sum_m u_m \left\{ \begin{matrix} ik \\ m \end{matrix} \right\} = \frac{1}{(n-1)!} \beta f'_{ik} (f) \sum_m u_m F^{1,m},$$

which gives, with a slight change of notation,

$$\sum_m u_m \left\{ \begin{matrix} ik \\ m \end{matrix} \right\} = \frac{1}{(n-1)!} f_{ik} (ua) (fa). \quad (49)$$

Now Bianchi's covariant second derivatives of  $u$  are defined as follows:<sup>12</sup>

$$u_{(ik)} = \frac{\partial^2 u}{\partial x_i \partial x_k} - \sum_m \frac{\partial u}{\partial x_m} \left\{ \begin{matrix} ik \\ m \end{matrix} \right\}. \quad (50)$$

Hence

$$(n-1)! u_{(ik)} = (n-1)! u_{ik} - f_{ik} (ua) (fa),$$

and by means of (43)

$$(n-1)! u_{(ik)} = f_i (ua)_k (fa), \quad (51)$$

which formula verifies the above statement concerning the coefficients of the quadratic covariant expression (46).

This formula leads now also to the proof of (28).

Bianchi defines<sup>13</sup>

$$\Delta_2 u = \beta^2 \sum_{ik} A_{ik} u_{(ik)}, \quad (52)$$

which gives

$$(n-1)! \Delta_2 u = \beta^2 (fa) \sum_{ik} A_{ik} f_i (ua)_k.$$

But from (21)

$$\beta^2 \sum_{ik} A_{ik} f_i (ua)_k = \frac{1}{(n-1)!} (f\phi) ((ua), \phi),$$

and from (34)

$$(fa) (f\phi) ((ua), \phi) = (n-1)! ((ua), a),$$

therefore

$$(n-1)! \Delta_2 u = ((ua), a).$$

<sup>11</sup> CHRISTOFFEL, *loc. cit.*, p. 49, and BIANCHI, *loc. cit.*, p. 43.

<sup>12</sup> *Loc. cit.*, p. 46.

<sup>13</sup> *Loc. cit.*, p. 47, formula (24).

## §7. THE QUADRILINEAR COVARIANT AND THE SIMPLEST INVARIANT

In the preceding article we found a quadratic covariant expression. It can be shown that the lowest covariant proper, *i. e.*, one which does not contain any arbitrary function  $u$ , is quadrilinear and of the second order. This covariant occurs in Riemann's paper: "Commentatio mathematica, qua respondere tentatur quaestioni ab ill<sup>ma</sup> Accademia Parisiensi propositae: Trouver quel doit être l'état calforique d'un corps solide homogène indéfini pour qu'un système de courbes isothermes, à un distant donné, restent isothermes après un temps quelconque, de telle sorte que la température d'un point puisse s'exprimer en fonction du temps et de deux autres variables indépendantes." The second part of this paper is devoted practically to the analytic deduction of the propositions established in the famous paper: "Ueber die Hypothesen welche der Geometrie zu Grunde liegen." The quadrilinear covariant constitutes the numerator of a fraction which Riemann defines as the general measure of curvature.<sup>14</sup>

Christoffel arrived quite independently<sup>15</sup> at the same covariant; it forms the basis of his deductions of the conditions for the equivalence of two quadratic differential quantics.

The covariant is defined in the four sets of differentials,  $d^{(1)}x, d^{(2)}x, d^{(3)}x, d^{(4)}x$  as

$$G_i = \sum_{ikrs} (ikrs) d^{(1)}x_i d^{(2)}x_k d^{(3)}x_r d^{(4)}x_s, \quad (53)$$

where

$$(ikrs) = \frac{\partial \left[ \begin{smallmatrix} ir \\ k \end{smallmatrix} \right]}{\partial x_s} - \frac{\partial \left[ \begin{smallmatrix} is \\ k \end{smallmatrix} \right]}{\partial x_r} + \beta^2 \sum_{lm} A_{lm} \left\{ \left[ \begin{smallmatrix} is \\ m \end{smallmatrix} \right] \left[ \begin{smallmatrix} rk \\ l \end{smallmatrix} \right] - \left[ \begin{smallmatrix} ir \\ m \end{smallmatrix} \right] \left[ \begin{smallmatrix} sk \\ l \end{smallmatrix} \right] \right\}. \quad (54)$$

If now we introduce symbols, we have

$$(ikrs) = f_{ir} f_{ks} - f_{is} f_{kr} + \frac{\beta^2}{(n-1)!} \sum_{lm} F^{1,l} \cdot F^{1,m} \cdot [f_m f_{is} \phi_l \phi_{rk} - f_m f_{ir} \phi_l \phi_{sk}],$$

which expression, considering that

$$\beta \cdot \sum_i \phi_i F^{1,i} = (\phi f), \quad \beta \cdot \sum_m f_m F^{1,m} = (f),$$

is transformed into

$$(ikrs) = f_{ir} f_{ks} - f_{is} f_{kr} + \frac{1}{(n-1)!} (f_{is} \phi_{rk} - f_{ir} \phi_{sk}) (\phi f) (f). \quad (55)$$

But since now, according to (43)

$$\begin{aligned} \phi_{rk} (\phi f) (f) &= (n-1)! f_{kr} - \phi_k (f)_r (\phi f), \\ \phi_{sk} (\phi f) (f) &= (n-1)! f_{ks} - \phi_k (f)_s (\phi f), \end{aligned}$$

<sup>14</sup> Riemann's *gesammelte Werke*, herausgegeben von H. WEBER, 2d ed. (Leipzig, 1892), pp. 403, 412.

<sup>15</sup> In 1869. RIEMANN's paper "Commentatio mathematica" was written 1861, but published only ten years after his death by Weber-Dedekind.

<sup>16</sup> CHRISTOFFEL, *loc. cit.*, p. 53; BIANCHI, *loc. cit.*, p. 50.

<sup>17</sup> RIEMANN, *loc. cit.*, pp. 402, 411; CHRISTOFFEL, *loc. cit.*, p. 54; BIANCHI, *loc. cit.*, p. 51.



the first terms in (55) cancel out so that after a slight change of notation we have

$$(n-1)! (ikrs) = [f_{ir}(fa)_s - f_{is}(fa)_r] (\phi a) \phi_k . \quad (56)$$

This leads by means of

$$f_{ir}(fa)_s (\phi a) = \frac{1}{n} (fa)_r (fa)_s \phi_i - f_i (\phi a)_r (fa)_s ,$$

and a similar relation with  $r$  and  $s$  interchanged which follows immediately from (44) to

$$(n-1)! (ikrs) = f_i \phi_k [(fa)_r (\phi a)_s - (fa)_s (\phi a)_r] . \quad (57)$$

Multiplying out and interchanging in the second term the two equivalent symbols  $f$  and  $\phi$ , we have also

$$(n-1)! (ikrs) = (fa)_r (\phi a)_s [f_i \phi_k - f_k \phi_i] , \quad (58)$$

and interchanging  $f$  and  $\phi$  in both terms of (57) or (58)

$$(n-1)! (ikrs) = \frac{1}{2} [(fa)_r (\phi a)_s - (fa)_s (\phi a)_r] \cdot [f_i \phi_k - f_k \phi_i] . \quad (59)$$

To each of the parentheses appearing in (57), (58), (59) we now apply formula (37), and thus we find, changing the notation of the symbols, the final result

$$(n-1)! (n-2)! (ikrs) = ((af), (bf), c) (c) a_i b_k c_r^i c_s^k , \quad (60)$$

$$(n-1)! (n-2)! (ikrs) = (f\phi c) (c) c_i^i c_k^k (fa)_r (\phi a)_s , \quad (61)$$

$$2(n-1)! (n-2)! (n-2)! (ikrs) = ((fc), (\phi c), b) (f\phi a) (a) (b) a_i^i a_k^k b_r^i b_s^k . \quad (62)$$

Each one of these three formulas leads to a symbolic representation of  $G_4$ , which at the same time furnishes the proof that  $G_4$  is a covariant of  $A$ . From (60), for instance, we have

$$(n-1)! (n-2)! G_4 = ((af), (bf), c) (c) a_{2^{(n)}} b_{2^{(n)}} c_{2^{(n)}}^i c_{2^{(n)}}^k . \quad (63)$$

Closely related with this covariant  $G_4$  is the following *invariant of order two*:

$$I_2 = ((fa), (\phi a), c) (f\phi c) . \quad (64)$$

As its form shows, it is an invariant for every  $n$  and represents in particular for  $n = 2$  the Gaussian curvature:

$$((fa), (\phi a)) (f\phi) = 2K .$$

**SUBGROUPS OF THE GENERALIZED FINITE  
MODULAR GROUP**



# THE SUBGROUPS OF THE GENERALIZED FINITE MODULAR GROUP

ELIAKIM HASTINGS MOORE

## INTRODUCTION

THE modular group  $\Gamma$  of all unimodular substitutions  $(a, \beta; \gamma, \delta)$

$$\omega' = \frac{a\omega + \beta}{\gamma\omega + \delta} \quad (a\delta - \beta\gamma = 1)$$

of the complex variable  $\omega$ , where the  $a, \beta, \gamma, \delta$  are rational integers, has for every rational prime  $q$  a self-conjugate subgroup  $\Gamma_{\mu(q)}$  of finite index  $\mu(q)$  containing all substitutions  $(a, \beta; \gamma, \delta)$  for which  $a \equiv 1, \beta \equiv 0, \gamma \equiv 0, \delta \equiv 1 \pmod{q}$ . The corresponding quotient-group  $\Gamma/\Gamma_{\mu(q)}$  is conveniently given as the say finite modular group  $G_{\mu(q)}^{q^n+1}$  of substitutions  $(a, \beta; \gamma, \delta)$  on the  $q+1$  marks  $\omega$  ( $\omega = \infty, 0, 1, \dots, q-1$ ) where the  $a, \beta, \gamma, \delta$  are integers taken modulo  $q$ . (Cf. Klein-Fricke, *Theorie der elliptischen Modul-functionen*, Vol. I, pp. 419-91.) By generalizing from the Galois-field of rank 1 to that of rank  $n$  ( $GF[q^n]$ ; Galois's imaginaries) we have the generalized finite modular group  $G_{M(q^n)}^{q^n+1}$  of order  $M(q^n) = q^n(q^{2n}-1)$  or  $q^n(q^{2n}-1)/2$  according as  $q=2$  or  $q>2$ . Mathieu first exhibited this group and studied its cyclic subgroups (*Comptes rendus*, 1858, 1859; *Liouville's Journal*, 1860, p. 39). That (except for the cases  $q^n=2^1, 3^1$ ) it is simple was proved by myself (*Mathematical Papers of the Chicago Congress*, 1893, pp. 208-42; published in 1896; in abstract, *Bulletin of the New York Mathematical Society*, December, 1893, Vol. III, pp. 73-8), and independently, but incompletely, by Burnside (*Proceedings of the London Mathematical Society*, Vol. XXV, pp. 113-39, February, 1894).

In the present paper all subgroups of the  $G_{M(q^n)}^{q^n+1}$  are determined; for the case  $n=1$  this was done by Gierster (*Mathematische Annalen*, Vol. XVIII (1881)). The subgroups are of three kinds: (1) metacyclic or solvable groups; (2) (3) groups of the abstract character of certain groups  $G_{M(q^n)}^{q^n+1}$  or of certain groups  $G_{sM(q^n)}^{q^n+1}$  ( $q>2$ ), a group  $G_{sM(q^n)}^{q^n+1}$  being obtained by extending the  $G_{M(q^n)}^{q^n+1}$  by the substitution  $\omega' = \rho\omega$ , where  $\rho$  is a primitive root of the  $GF[q^n]$ . Thus the doubly infinite system of simple groups

$$G_{M(q^n)}^{q^n+1} \quad (q \neq 2^1, 3^1)$$

determines by the decomposition of the subgroups of its constituent groups, apart from the simple groups of prime order, only simple groups of the original system. An equation of degree  $q^n+1$  ( $q^n \neq 2^1, 3^1$ ) whose Galois-group is the  $G_{M(q^n)}^{q^n+1}$  has resolvents

of degree  $D$ ,  $D < q^n + 1$ , only in the cases  $q^n = 5^1, 7^1, 11^1, 3^3$ , when  $D$  is respectively 5, 7, 11, 6. For  $n = 1$  this is a noted theorem of Galois and Gierster.

The group  $\Gamma[\Omega]$  of all unimodular substitutions  $(\alpha, \beta; \gamma, \delta)$  where the coefficients are algebraic integers of an algebraic field  $\Omega$  [Körper  $\Omega$ ] has for every prime functional  $\pi$  [Weber] of absolute norm  $q^n$ , a self-conjugate subgroup  $\Gamma_{M(q^n)}[\Omega]$ . The quotient-group is the group  $G_{M(q^n)}^{q^n+1}$ . The substitutions of the modular group  $\Gamma$  leave invariant a modular function  $J(\omega)$ . The transcendental modular equation  $J = J(\omega)$  has algebraic resolvents, modular equations, with Galois-groups  $G_{M(q)=\mu(q)}^{q+1}$ . The group  $\Gamma[\Omega]$ , for field  $\Omega$  of degree  $> 1$ , is improperly discontinuous and has no corresponding automorphic function. If however an automorphic function  $K(\omega)$  belongs to a group whose substitutions have coefficients of an algebraic field  $\Omega$ , then the generalized modular equation  $K = K(\omega)$  has algebraic resolvents whose Galois-groups are certain generalized finite modular groups  $G_{M(q^n)}^{q^n+1}$  or certain of their subgroups. This remark connects the present paper with the recent work of Fricke and Bianchi on certain special classes of automorphic functions (cf. Fricke-Klein, *Automorphe Functionen*, Vol. I (1897), pp. 585 ff.).

This paper was read before the American Mathematical Society at the Boston meeting, August, 1898. The preceding paragraphs are, with insignificant modifications, the abstract of the paper incorporated (pp. 7, 8) in the report of the secretary of the society as published in the *Bulletin* of the society for October, 1898 (Vol. V, No. 1). The paper was to be offered to the *Mathematische Annalen*. Wishing, however, to develop the ideas of the last paragraph of the abstract, I held the paper back. The following year, April 12, 1899, Wiman presented to the Swedish Academy his memoir: *Bestimmung aller Untergruppen einer doppelt unendlichen Reihe von einfachen Gruppen* (Bihang till K. Svenska Vet.-Akad. Handlingar, Vol. XXV, Part I, No. 2).

Only slightly modified from its form of August, 1898, I am now publishing my paper. I enter into no comparison with the treatments of Burnside (*loc. cit.*) for  $q = 2$ , of Wiman and of Dickson (*Linear Groups*, Teubner, 1901, §§ 68–71, 239–63, based on my MS. and the memoir of Wiman).

In § 1 the theorems of generational determination of the symmetric and the alternating groups are stated. In §§ 2, 3, the properties of the finite field (endliche Körper), in particular that every finite field is the abstract form of a Galois-field, are developed. §§ 4, 5 are devoted to the definitions of the “real” and two “imaginary” forms of the group under consideration, while in § 6 the individual operators and the “largest” cyclic and commutative subgroups of the  $G_{M(s)} (s = q^n)$  are determined and classified as to conjugacy. §§ 2–6 come for the most part without change from my Chicago Congress paper of 1893, and form an integral part of this paper. In § 8 is proved the simplicity of the  $G_{M(s)} (s \neq 2^1, 3^1)$ ; in 1893 I proved this theorem by the elaborate consideration of a single complicated diophantine equation; in the present

proof two of three cases are treated by the use of theorems stated by Burnside and proved in § 3, 51°, and § 7, and the third case is treated by a simple diophantine equation. Of course, the complete enumeration of the subgroups determines this question of simplicity; but a more direct proof is clearly desirable.

In the complete enumeration the new "second imaginary" form  $G_{M(q)}^{s'+1}$  (§ 5 (15)) of the  $G_{M(q)}$  proves to be serviceable. The study in § 10 of the conjugacy of the commutative subgroups of order a power of  $q$  reduces to the study of the multiplier Galois-fields of the additive groups of the basal Galois-field of order  $q^s$ , and in this connection in § 11 we meet interesting relations of cyclotomic functions.

The theorem of § 17, generalizing the theorem of Galois and Gierster, as to the exceptional cases  $s = q^s = 5^1, 7^1, 11^1, 3^2$ , I communicated to Mr. Fricke in a letter dated February 17, 1894.

### §1. THE ABSTRACT SYMMETRIC AND ALTERNATING GROUPS

(1°) We have the important twofold:

*Theorem I.*—If the generators  $\left\{ \begin{matrix} B_0, \dots, B_{k-2} \\ E_1, \dots, E_{k-2} \end{matrix} \right\}$  of a group  $\left\{ \begin{matrix} G\{B_0, \dots, B_{k-2}\} \\ G\{E_1, \dots, E_{k-2}\} \end{matrix} \right\}$  of order greater than  $\left\{ \begin{matrix} 2; 6 \\ 1; 3 \end{matrix} \right\}$  for  $\left\{ \begin{matrix} k \neq 4; k=4 \\ k \neq 4; k=4 \end{matrix} \right\}$  satisfy the equations

$$\left\{ \begin{matrix} I = B_d^2 = (B_d B_{d+1})^2 = (B_d B_e)^2 & \left( \begin{matrix} d, d+1, e = 0, 1, \dots, k-2 \\ e > d+1 \end{matrix} \right) \end{matrix} \right. \quad (1)$$

$$\left\{ \begin{matrix} I = E_1^2 = E_d^2 = (E_d E_{d+1})^2 = (E_d E_e)^2 & \left( \begin{matrix} d', d, d+1, e = 1, 2, \dots, k-2 \\ d' > 1, e > d+1 \end{matrix} \right) \end{matrix} \right. \quad (2)$$

then the group  $G$  is of order  $\left\{ \begin{matrix} k! \\ \frac{1}{2} k! \end{matrix} \right\}$  and is holodrically isomorphic with the  $\left\{ \begin{matrix} \text{symmetric} \\ \text{alternating} \end{matrix} \right\}$  group on  $k$  letters, and the generators satisfy no further independent relations.

In the substitution-groups on the  $k$  letters  $l_1, \dots, l_k$  the generators  $B_d, E_d$  are, for instance,

$$B_d = (l_{d+1} l_{d+2}), \quad E_d = B_0 B_d = (l_1 l_2)(l_{d+1} l_{d+2}) \quad (d, d+1 = 2, 3, \dots, k). \quad (3)$$

We denote the abstract symmetric and alternating groups by the notations  $G_{k!}^{(k)}, G_{\frac{1}{2}k!}^{(k)}$ .

(2°) *Theorem II.*—An abstract group  $G$  of elements  $\dots P \dots$  having the schematic multiplication table

$$P_i = P_1 P_i,$$

the transposed multiplication table

$$Q_i = Q_i Q_1$$

in the elements  $\dots Q \dots$  in 1-1 notational correspondence with the elements  $\dots P \dots$  defines an abstract group  $\tilde{G}$ —the transpose group of the original group.

<sup>1</sup> MOORE, "Concerning the Abstract Groups of Order  $k!$  and  $\frac{1}{2}k!$  Holodrically Isomorphic with the Symmetric and Alternating Substitution-Groups on  $k$  Letters," *Proceedings of the London Mathematical Society*, December 10, 1896, Vol. XXVIII (1897), pp. 357-66; Theorems A, B, and C.

A group  $G$  identical with its transpose  $\bar{G}$  is called a self-transpose group.

(3°) *Theorem III.*—The abstract symmetric and alternating groups  $G_{k!}^{(k)}$ ,  $G_{\frac{1}{2}k!}^{(k)}$  are self-transpose groups.—To prove for the  $G_{\frac{1}{2}k!}^{(k)}$  with elements  $P$  and generators  $E$ : The group  $G_{\frac{1}{2}k!}^{(k)} \{E_1, \dots, E_{k-2}\}$  is holodrically transformed into itself  $G_{\frac{1}{2}k!}^{(k)} \{F_1, \dots, F_{k-2}\}$ , where  $F_d = E_d^{-1}$  ( $d = 1, \dots, k-2$ ), since the  $F_d$  satisfy the relations (2) of the  $E_d$ . But in view of the formula  $(P_1 P_2)^{-1} = P_2^{-1} P_1^{-1}$  the two groups with elements  $\dots P \dots, \dots Q \dots$ , where  $Q = P^{-1}$  are seen to be transposes each of the other.

(4°) In the substitution-group  $G_{M(n)}^{s+1}$  of this paper (§ 4) it is, owing to the analytic form of the substitutions, convenient to denote the substitution compounded from two substitutions, first  $S$ , then  $T$ , by  $TS$  instead of, as usual, by  $ST$ .—In view of Theorem III this notational variation will cause no complication in the application of Theorem I.

(5°) *Examples* of concrete forms of the symmetric and alternating groups—in the notations<sup>2</sup> of § 4.

#### SYMMETRIC GROUPS $G_{k!}$

	$k$	$k!$	Group	$E_1$	$E_2$	$E_3$	$E_4$	$E_5$
(I)	3	6	$G_{M(3)}^3$	$\begin{pmatrix} 1, 1 \\ 0, 1 \end{pmatrix}$	$\begin{pmatrix} 1, 0 \\ 1, 1 \end{pmatrix}$			
(II)	4	24	$G_{M(3)}^4$	$\begin{pmatrix} -1, 0 \\ 0, 1 \end{pmatrix}$	$\begin{pmatrix} -1, -1 \\ 0, 1 \end{pmatrix}$	$\begin{pmatrix} 0, 1 \\ 1, 0 \end{pmatrix}$		
(III)	5	120	$G_{M(5)}^5$	$\begin{pmatrix} 1, -1 \\ 3, -1 \end{pmatrix}$	$\begin{pmatrix} 2, 1 \\ -1, 3 \end{pmatrix}$	$\begin{pmatrix} 0, -1 \\ 2, 0 \end{pmatrix}$	$\begin{pmatrix} 1, 2 \\ 3, -1 \end{pmatrix}$	
(IV)	6	720	$\bar{G}_{M(3^2)}^{10}$	$\begin{pmatrix} 1, 0 \\ 0, 1 \end{pmatrix}$	$\begin{pmatrix} 1, \sqrt{-1} \\ 0, 1 \end{pmatrix}$	$\begin{pmatrix} 0, 1 \\ 1, 0 \end{pmatrix}$	$\begin{pmatrix} -1, 1 \\ 0, 1 \end{pmatrix}$	$\begin{pmatrix} -1, 0 \\ 0, 1 \end{pmatrix}$

#### ALTERNATING GROUPS $G_{\frac{1}{2}k!}$

	$k$	$\frac{1}{2}k!$	Group	$E_1$	$E_2$	$E_3$	$E_4$
(V)	4	12	$G_{M(3)}^4$	$\begin{pmatrix} 1, 1 \\ 0, 1 \end{pmatrix}$	$\begin{pmatrix} 0, -1 \\ 1, 0 \end{pmatrix}$		
(VI)	5	60	$G_{M(3^2)}^5$	$\begin{pmatrix} \rho, 0 \\ 0, \rho^{-1} \end{pmatrix}$	$\begin{pmatrix} 1, 1 \\ 0, 1 \end{pmatrix}$	$\begin{pmatrix} 0, -1 \\ 1, 0 \end{pmatrix}$	$(\rho^2 + \rho + 1 = 0)$
(VII)	5	60	$G_{M(6)}^6$	$\begin{pmatrix} 1, 1 \\ -1, 0 \end{pmatrix}$	$\begin{pmatrix} 1, 3 \\ 1, 4 \end{pmatrix}$	$\begin{pmatrix} 1, 1 \\ 0, -1 \end{pmatrix}$	
(VIII)	6	360	$G_{M(3^2)}^{10}$	$\begin{pmatrix} 1, -\sqrt{-1} \\ 0, 1 \end{pmatrix}$	$\begin{pmatrix} 0, 1 \\ 1, 0 \end{pmatrix}$	$\begin{pmatrix} -1, 1 \\ 0, 1 \end{pmatrix}$	$\begin{pmatrix} -1, 0 \\ 0, 1 \end{pmatrix}$

<sup>2</sup>Except for the group (IV)  $\bar{G}_{M(3^2)}^{10}$ . In this case the coefficients are marks of the  $GF[3^2]$  and the substitution the conjugate of  $\omega$  with respect to the  $GF[3^2]$  (§ 3 (48°) § 4;  $\bar{\omega} = \omega$ ), is denoted by

$$\omega' = \frac{\alpha\bar{\omega} + \beta}{\gamma\bar{\omega} + \delta} \quad \left( \frac{\alpha, \beta}{\gamma, \delta} \right)_-$$

on the 10 marks  $\omega$ —the 9 of the  $GF[3^2]$  and  $\infty$ —where  $\bar{\omega}$  is or  $(\alpha, \beta; \gamma, \delta)_-$ . (The  $G_{M(3^2)}^{10}$  is not a symmetric  $G_{6!}$ .)

(6°) The regular tetrahedron, octahedron, icosahedron have groups  $G_{12}$ ,  $G_{24}$ ,  $G_{60}$  of rotation into themselves which are holoedrically isomorphic to the  $G_{12}^t$ ,  $G_{24}^t$ ,  $G_{60}^t$ . Accordingly these latter groups and all corresponding group-forms are called the tetrahedron, octahedron, icosahedron groups.

(7°) We notice (for use in § 15 (2°)) that the tetrahedron group  $G_{12} \{E_1, E_2\}$  [ $I = E_1^3 = E_2^3 = (E_1 E_2)^3$ ] contains a self-conjugate four-group [ $I, K = E_1, L = E_1 E_2 E_1^2, M = KL = E_1^2 E_2 E_1$ ] and an operator  $U = E_1$  such that

$$U^4 = I, \quad UK = LU, \quad UL = MU, \quad UM = KU.$$

## § 2. THE GALOIS-FIELD OF ORDER $s = q^n$ , $GF[s]$

Suppose that we have a system of  $s$  distinct symbols or *marks*,<sup>3</sup>  $\mu_1, \mu_2, \dots, \mu_s$  ( $s$  being some finite positive integer), and suppose that these marks may be combined by the four fundamental operations of algebra—addition, subtraction, multiplication, and division—these operations being subject to the ordinary abstract *operational identities* of algebra

$$\mu_i + \mu_j = \mu_j + \mu_i; \quad \mu_i \mu_j = \mu_j \mu_i; \quad (\mu_i + \mu_j) \mu_k = \mu_i \mu_k + \mu_j \mu_k; \quad \text{etc.}$$

and that when the marks are so combined the results of these operations are in every case uniquely determined and belong to the system of marks. Such a system of marks we shall call a *field of order  $s$* , using the notation  $F[s]$ .

We are led at once to seek *To determine all such fields  $F[s]$  of order  $s$ .*

This determination is the subject of this section, § 2.

The most familiar instance of such a field of order  $s = q$  = a prime is the system of  $q$  incongruous classes (modulo  $q$ ) of rational integral numbers  $a$ .

Galois discovered an important generalization<sup>4</sup> of the preceding field. Let  $F_n(X)$ , a rational integral function of the indeterminate  $X$  of degree  $n$  with integral coefficients,  $c_i$ ,  $c_n = +1$ ,

$$F_n(X) \equiv \sum_{i=0}^n c_i X^i,$$

be irreducible, modulo  $q$ . Then the Galois-field of order  $s = q^n$ ,  $GF[q^n]$ , consists of the system of  $q^n$  incongruous classes (modulus  $q$ ,  $F_n(X)$ ) of rational integral functions of  $X$  with integral coefficients. In this  $GF[q^n]$  there exist *primitive roots*; the  $q^n - 1$  successive powers of a primitive root are the  $q^n - 1$  marks of the field (0 excluded).

*The Galois-field  $GF[q^n]$  is uniquely defined for every  $q$  = prime,  $n$  = positive integer; that is:*

*$F_n(X)$  which are irreducible (mod.  $q$ ) do exist;*

*The  $GF[q^n]$  is independent of the particular irreducible  $F_n(X)$  used in its construction.*

<sup>3</sup>It is necessary that all quantitative ideas should be excluded from the concept marks. Note that the signs  $>$ ,  $<$  do not occur in the theory.

<sup>4</sup>For the details of this Galois theory, see GALOI "Sur la théorie des nombres," *Bulletin des sciences math.*

*matiques de M. Férussac*, Vol. XIII (1830), p. 428; reprinted in *Journal de mathématiques pures et appliquées*, Vol. XI (1846), pp. 396-407; SERRIN, *Algèbre supérieure*, fifth edition, Vol. II, pp. 122-69; and JORDAN, *Substitutions*, pp. 14-18.



Assuming now (nothing but) the existence of a field of order  $s$   $F[s]$ , I proceed to establish its fundamental properties, and prove in particular (40°) that:

*Every existent field  $F[s]$  is the abstract form of a Galois-field  $GF[q^n]$ ;  $s = q^n$ .*

The purely abstract form here given to the theory would seem to fit it best for immediate use wherever it can with advantage be introduced.

Naturally in many details my deduction of the properties of the  $F[s]$  runs parallel to the work of Galois, Serret, and Jordan in investigating the  $GF[q^n]$ . I forbear to give closer references. For *ultimate existence-proofs* I fall back (22°, 38°, 39°) on the Galois theory. This sharp separation of the necessary properties of the field  $F[s]$  if existent from the details of the various existence-proofs I consider highly desirable.

*Def.*—Any rational function of any number of indeterminates  $X_1, X_2, \dots, X_n$ , is said to *belong to the field* or to *be of the field*, if it has as coefficients marks of the field. An equation between two such functions is said likewise to belong to the field or to be of the field.

A rational integral function of any number of indeterminates belonging to the field is called *irreducible in the field* when it is not identically\* the product of two or more such functions belonging to the field.

(1°) The theorems of ordinary algebra concerning rational functions of indeterminates hold for functions of our field; in particular, the algorithm for determination of the highest common factor of two rational integral functions of the indeterminate  $X$ .

(2°) The  $s$  marks  $\mu_1 - \mu_1, \mu_2 - \mu_2, \dots, \mu_s - \mu_s$  are equal; this mark is written  $\mu_{(0)}$  or 0.

(3°) In division the mark  $\mu_{(0)} = 0$  may never be the divisor.

(4°) The  $s-1$  marks  $\frac{\mu_i}{\mu_i} (\mu_i \neq \mu_{(0)})$  are equal; this mark is written  $\mu_{(1)}$  or 1.

(5°)  $\mu_i + \mu_{(0)} = \mu_i$  for every  $\mu_i$ .

(6°)  $\mu_{(0)} \mu_i = \mu_i \mu_{(0)} = \mu_{(0)}$  for every  $\mu_i$ .

(7°)  $\mu_{(1)} \mu_i = \mu_i \mu_{(1)} = \mu_i$  for every  $\mu_i$ .

(8°)  $\mu_i \mu_j = \mu_{(0)} = 0$  only if  $\mu_i = \mu_{(0)} = 0$  or  $\mu_j = \mu_{(0)} = 0$ .

*Def.*— $c$  being any positive integer, we denote the mark  $\mu_{(1)} + \mu_{(1)} + \dots + \mu_{(1)}$  or  $1 + 1 + \dots + 1$  ( $c$  terms) by  $\mu_{(c)}$  or  $c$  (but by the latter notation  $c$  only when it is perfectly clear that by  $c$  a mark  $c$  is meant).

$$\mu_{(0)} - \mu_{(1)} = \mu_{(-1)} .$$

$$\mu_{(-1)} + \mu_{(-1)} + \dots + \mu_{(-1)} (c \text{ terms}) = \mu_{(-c)} .$$

Thus we have defined the marks for *all* integral values  $c$ . These are called the *integral marks* of the  $F[s]$ . (See footnote at beginning of § 2, and (17°).)

(9°)  $\mu_{(a_1)} + \mu_{(a_2)} = \mu_{(a_1+a_2)}$ ;  $\mu_{(a_1)} - \mu_{(a_2)} = \mu_{(a_1-a_2)}$ ;  $\mu_{(a_1)} \mu_{(a_2)} = \mu_{(a_1 a_2)}$ .

(10°)  $\mu_i + \mu_i + \dots + \mu_i$  ( $c$  terms)  $= \mu_{(c)} \mu_i = c \mu_i$ .

\* In the indeterminates.

(11°) The equation belonging to the field

$$f_k(X) = 0,$$

where  $f_k(X)$  is a rational integral function of  $X$  of degree  $k$ , has in the field at most  $k$  roots, unless it is an identity in  $X$  when every mark of the field is a root.

(12°) If we have in the field an identity in  $X$

$$f_k(X) \equiv f_{l_1}(X) \cdot f_{l_2}(X) \quad (k=l_1+l_2),$$

and if the equations  $f_k(X) = 0, f_{l_1}(X) = 0$  have in the field  $k'$  and  $l_1'$  roots respectively, then the equation  $f_{l_2}(X) = 0$  must have as roots the remaining  $k' - l_1'$  roots of  $f_k(X) = 0$ .

Here

$$k' \leq k, l_1' \leq l_1, k' - l_1' \leq k - l_1.$$

In particular, if  $k' = k$ , then  $l_1' = l_1$ , for from  $l_1' < l_1$  would follow  $k' - l_1' > k - l_1$ .

(13°) For every mark  $\nu$  of the field  $F[s]$  there exist (because there are only  $s$  marks in the field) positive integral solutions  $c$  of the equation

$$c\nu = 0, \text{ or } \mu_{(c)}\nu = 0.$$

The smallest such solution is called the additive-period of the mark  $\nu$ . All the solutions are multiples of this additive-period.

(14°) The additive-period of the mark  $\mu_{(0)} = 0$  is 1.

(15°) The additive-period  $c_\nu$  of any mark  $\nu \neq \mu_{(0)}$  is the smallest positive integer  $c$  for which (7°, 13°)

$$\mu_{(c)} = \mu_{(c)}\mu_{(1)} = \mu_{(c)},$$

that is, is the additive-period  $c_{\mu_{(1)}}$  of the mark  $\mu_{(1)}$  or 1. All marks  $\nu \neq \mu_{(0)}$  have the same additive-period  $c$ . This integer  $c$  is called the modulus of the field  $F[s]$ .

(16°) This common additive-period or modulus  $c$  is a prime  $q$ . For if  $c$  were composite,  $c = dd'$  ( $d < c, d' < c$ ), we should have (9°, 13°)

$$\mu_{(c)} = \mu_{(d)}, \mu_{(c)} = \mu_{(d)}\mu_{(d')} = \mu_{(d)}, \mu_{(d)} \neq \mu_{(c)},$$

whence  $\mu_{(d')} = \mu_{(0)}$ , while from  $d' < c$  follows  $\mu_{(d')} \neq \mu_{(0)}$ .

(17°) The marks  $c = \mu_{(c)}$  for  $c = 0, 1, 2, \dots, (q-1)$  are distinct; the mark  $q = \mu_{(q)} = \mu_{(0)} = 0$ . The integral marks are thus "to be taken modulo  $q$ ." (This inheres in the concept marks and is not indicated in the notation.) These  $q$  integral marks form a field  $F[q]$  (the abstract form of the well-known field of rational integers taken modulo  $q$ ).

(18°)  $\mu$  being any mark  $\neq \mu_{(0)}$ , the  $q$  marks  $c\mu$  ( $c = 0, 1, 2, \dots, (q-1)$ ) are distinct and form an additive-group  $[\mu]$  with the basis-system  $\mu$ , in the sense that the sum of any two of these marks is a third of the same system

$$c_1\mu + c_2\mu = (c_1 + c_2)\mu,$$

or

$$\mu_{(c_1)}\mu + \mu_{(c_2)}\mu = \mu_{(c_1+c_2)}\mu.$$

*Def.*—Any  $h$  marks  $\nu_i$  ( $i = 1, 2, \dots, h$ ) of the field  $F[s]$  are called linearly independent with respect to the field  $F[q]$  (17°) if the equation

$$\sum_{i=1}^h c_i \nu_i = 0,$$

where the  $c_i$ 's are marks of the  $F[q]$ , can be satisfied only in case every  $c_i = 0$ .

(19°) Any  $h$  marks  $\nu_i$  ( $i = 1, 2, \dots, h$ ) linearly independent with respect to the  $F[q]$  give rise to the  $q^h$  distinct marks of the field

$$\sum_{i=1}^h c_i \nu_i. \quad (\text{Every } c_i \text{ has the values } 0, 1, \dots, (q-1)).$$

*Def.*—These  $q^h$  marks form an additive-group  $[\nu_1, \nu_2, \dots, \nu_h]$  of rank<sup>6</sup>  $h$  with respect to the field  $F[q]$  of which the  $h$  marks  $\nu_1, \dots, \nu_h$  form the basis-system.<sup>6</sup> The additive-group is transformed into itself when every mark is multiplied by a mark ( $\neq 0$ ) of the  $F[q]$ .

If from any  $h'$  linearly independent marks  $\nu_j$ , ( $j = 1, 2, \dots, h'$ ) of this additive-group we form the additive-group  $[\nu'_1, \nu'_2, \dots, \nu'_h]$  of rank  $h'$ , this group is entirely contained in the original additive-group. Any  $h+1$  marks of an additive-group of rank  $h$  are linearly dependent.

Any mark of the field  $F[s]$  not in the additive-group  $[\nu_1, \nu_2, \dots, \nu_h]$  of rank  $h$  forms with the  $h$  marks of the basis-system a system of  $h+1$  linearly independent marks, the basis-system of an additive-group of rank  $h+1$ .

(20°) Our field  $F[s]$  of finite order  $s$  may therefore be exhibited as an additive-group of some finite rank  $n$ ; hence  $s$  must have the form  $s = q^n$ .

(21°) Any field included within our field  $F[s = q^n]$  may likewise be exhibited as an additive-group of rank  $l \leq n$ .

This rank  $l$  of the  $F[q^l]$  thought of as an additive-group is called also the rank of the field  $F[q^l]$  itself. It turns out (26°) that the rank  $l$  of the included field  $F[q^l]$  is a divisor of the rank  $n$  of the including field  $F[q^n]$ .

(22°) Any mark  $\mu$  of the field  $F[s = q^n]$  satisfies an equation of the form

$$f_k(X) \equiv \sum_{i=0}^k c_i X^i = 0 \quad (c_k = 1),$$

where the expression  $f_k(X)$  belongs to and is irreducible in the field  $F[q^l]$ . The positive integer  $k$ , the degree of the equation, is called the rank of the mark  $\mu$  with respect to the  $F[q^l]$ .

For consider the successive powers  $\mu^i$  ( $i = 0, 1, 2, \dots$ ) of the  $\mu$ . A linear relation with coefficients belonging to the  $F[q^l]$  holds certainly between any  $n+1$  marks of the  $F[s = q^n]$  (20°, 19°) and so between the first  $n+1$  powers  $\mu^i$ . Suppose that for our mark  $\mu$  such a linear relation does not hold between the first  $k$  powers (and so not between the first  $k'$  powers  $k' \leq k$ ), but does hold between the first  $k+1$  powers.

<sup>6</sup> FROBENIUS UND STICKELBERGER, "Ueber Gruppen von vertauschbaren Elementen," *Journal für die reine und angewandte Mathematik*, Vol. LXXXVI (1878), pp. 217-62, use these terms (p. 219).

We have

$$\sum_{i=0}^k c'_i \mu^i = 0 ,$$

where the  $c'_i$  belong to the  $F[q]$  and  $c'_k \neq 0$ ; divide by  $c'_k$  and set  $c_i$  for  $c'_i/c'_k$ ; we have

$$\sum_{i=0}^k c_i \mu^i = 0 ,$$

where the  $c_i$  belong to the  $F[q]$  and  $c_k = 1$ ; whence the equation

$$f_k(X) \equiv \sum_{i=0}^k c_i X^i = 0$$

is satisfied by the mark  $X = \mu$ .

The expression  $f_k(X)$  so determined does belong to the  $F[q]$  and is in it irreducible.<sup>1</sup> For if we had the decomposition  $f_k(X) \equiv f_{k_1}(X) f_{k_2}(X)$  in the field  $F[q]$ , then  $X = \mu$  would satisfy either  $f_{k_1}(X) = 0$  or  $f_{k_2}(X) = 0$ , neither of which equations can be satisfied by  $\mu$ , since the first  $k_1 + 1$ ,  $k_2 + 1 < k + 1$  powers of  $\mu$  are linearly independent with respect to the  $F[q]$ .

Following Galois (see the reference given above), we recognize in this *additive-group*  $[\mu^0 = 1, \mu^1, \dots, \mu^{k-1}]$  of rank  $k$  the abstract form of a Galois-field of order  $q^k$ ,  $GF[q^k]$ .

(23°) Every mark of our field  $F[s=q^k]$  serves to define such a Galois-field, the field of lowest rank  $k$  in which it lies. Every field in which it lies has its rank  $k'$  a multiple of  $k$  (26°).

Def.—Any  $h$  marks  $\nu_i$  ( $i = 1, 2, \dots, h$ ) of the field  $F[q^n]$  are called linearly independent with respect to any included field  $F[q^l]$  ( $l \leq n$ ), if the equation

$$\sum_{i=1}^h \gamma_i \nu_i = 0 ,$$

where the  $\gamma_i$  are marks of the field  $F[q^l]$ , can be satisfied only in case every  $\gamma_i = 0$ .

(24°) Any such system of  $h$  marks gives rise to  $q^h$  distinct marks of the field obtained from the general form

$$\sum_{i=1}^h \gamma_i \nu_i$$

by letting the  $h$   $\gamma_i$ 's run independently through the series of marks of the  $F[q^l]$ . These  $q^h$  marks form an *additive-group*  $[\nu_1, \nu_2, \dots, \nu_h | F[q^l]]$  with the *basis-system*  $\nu_1, \nu_2, \dots, \nu_h$  and the *field of reference*  $F[q^l]$ , of rank  $h$  with respect to the  $F[q^l]$ .

(25°) Any mark  $\nu_{h+1}$  of the field  $F[q^n]$  but not in this additive-group (24°) forms with the  $h$  marks of the basis-system a system of  $h + 1$  marks linearly independent with respect to the  $F[q^l]$ , and thus leads to an additive-group of rank  $h + 1$  with respect to the  $F[q^l]$ .

For the equation

$$\sum_{i=1}^{h+1} \gamma_i \nu_i = 0$$

is impossible, first, if  $\gamma_{h+1} = 0$ , since the  $h$  marks  $\nu_1, \nu_2, \dots, \nu_h$  are by hypothesis

<sup>1</sup>  $f_k(X)$  is reducible in the field  $F[s]$ ; it has the factor  $X - \mu$ .

linearly independent with respect to the  $F[q']$ , and, *secondly*, if  $\gamma_{h+1} \neq 0$ , since this equation would lead (by multiplying by  $-1/\gamma_{h+1}$ , and by substituting for the marks  $-\gamma/\gamma_{h+1}$  which belong to the field  $F[q']$  the marks  $\gamma'_i$  respectively) to the equation

$$\nu_{h+1} = \sum_{i=1}^h \gamma'_i \nu_i ,$$

which would contradict the hypothesis that the  $\nu_{h+1}$  is not in the additive-group  $[\nu_1, \nu_2, \dots, \nu_h | F[q']]$ .

(26°) Our field  $F[s = q^n]$  of *finite* order  $s = q^n$  may be exhibited with respect to any included field  $F[q']$  of order  $q'$  as an additive-group of some *finite* rank  $h$ ;  $s = q^n = q'^n$ ;  $n = hl$ .

The rank  $l$  of any field  $F[q']$  included in the field  $F[q^n]$  of rank  $n$  is a divisor of  $n$ . The quotient  $h$  is the rank of the including field  $F[q^n]$  with respect to the included field  $F[q']$ ;  $n = hl$ .

(27°) For every mark  $\nu \neq \mu_0$  of the field  $F[s = q^n]$  there exist (because there are only  $s$  marks in the field) positive integral solutions  $e$  of the equation

$$\nu^e = \mu_0 = 1 .$$

The smallest such solution, say  $e$ , is called the *multiplicative-period* or *exponent* of the mark  $\nu$ ;  $\nu$  is said to *belong to this exponent*  $e$ . The  $e$  marks  $\nu^0 = 1, \nu^1, \nu^2, \dots, \nu^{e-1}$  are distinct, and form a multiplicative-group.

(28°) If two marks  $\nu_1, \nu_2$  belong to exponents  $e_1, e_2$  respectively which are relatively prime, their product belongs to the exponent  $e_1 e_2$ , and the marks

$$\nu_1^{d_1} \nu_2^{d_2} \quad \left( \begin{matrix} d_1 = 0, 1, 2, \dots, (e_1 - 1) \\ d_2 = 0, 1, 2, \dots, (e_2 - 1) \end{matrix} \right)$$

are  $e_1 e_2$  distinct marks.

(29°) The  $s - 1 = q^n - 1$  marks  $\mu \neq 0$  may be thought of as the elements of a multiplicative-group of order  $s - 1$ . The *period* of the element  $\nu$  is its exponent  $e$ . Whence  $e$  is a divisor of  $s - 1 = q^n - 1$ .

The equation of the field  $F[s = q^n]$

$$X^s - 1 = 0 ,$$

where  $e$  is the exponent of some mark  $\nu$  has in the field at most  $e$  roots (11°) (not having the root  $X = 0 = \mu_0$ ), and in fact it has the  $e$  roots  $\nu^i$  ( $i = 0, 1, 2, \dots, (e-1)$ ).

(30°) The equation of the field

$$X^{s-1} - 1 = X^{e^{n-1}} - 1 = 0$$

is satisfied by every one of the  $s - 1$  marks  $\nu \neq 0$ , since the exponent of every mark  $\nu$  is a divisor of  $s - 1$ .

(31°) The equation of the field

$$X^s - X = X^{e^n} - X = 0$$

is satisfied by every mark  $\nu$  of the field  $F[s = q^n]$ . Hence the decomposition in the  $F[s = q^n]$

$$X^{q^n} - X \equiv \prod_{i=1}^n (X - \mu_i) .$$

This is the generalization for the  $F[q^n]$  of Fermat's Theorem for the  $F[q^1]$ .

(32°) Converse of (29°).  $f$  being any divisor of  $s - 1 = q^n - 1$  the equation

$$X^f - 1 = 0$$

has in the  $F[s = q^n]$  exactly  $f$  roots.

For let  $s - 1 = fg$ . We have the identity in  $X$  in the field  $F[q^1]$  and so in the field  $F[s = q^n]$ .

$$X^{s-1} - 1 \equiv (X^f - 1)(X^{g-1} + X^{g-2} + \dots + X^1 + X^0 + 1) ;$$

whence (30°, 12°) the desired conclusion follows.

(33°) In particular, if  $s - 1 = q^n - 1 = p_1^{h_1} p_2^{h_2} \dots p_k^{h_k}$ , where the  $p$ 's are the distinct prime factors of  $s - 1$ , the equation

$$X^{p_i^{h_i}} - 1 = 0$$

for  $i$  any integer of the series  $1, 2, \dots, k$  has in the field  $F[s = q^n]$   $p_i^{h_i}$  roots  $\nu$ . The exponent  $e$  of every one of these roots is a divisor of  $p_i^{h_i}$  and hence is of the form  $p_i^{h'_i}$  ( $h'_i \leq h_i$ ). Of these roots  $p_i^{h'_i-1}$  belong to exponents which are factors of  $p_i^{h_i-1}$ , namely, the  $p_i^{h'_i-1}$  roots of the equation

$$X^{p_i^{h'_i-1}} - 1 = 0 .$$

The remaining  $p_i^{h_i} - p_i^{h'_i-1} = p_i^{h'_i} \left(1 - \frac{1}{p_i}\right)$  roots belong to the exponent  $p_i^{h'_i}$  itself.

Whence, by multiplication of marks belonging to the various exponents  $p_i^{h'_i}$  for  $i = 1, 2, \dots, k$ , we obtain in all

$$\prod_{i=1}^k p_i^{h'_i} \left(1 - \frac{1}{p_i}\right) = \phi(s - 1) = \phi(q^n - 1)$$

distinct marks belonging by (28°) to the exponent

$$\prod_{i=1}^k p_i^{h'_i} = s - 1 = q^n - 1 .$$

[Here  $\phi(t)$  denotes, as usual, the number of integers less than and relatively prime to the positive integer  $t$ .]

Thus, the  $s - 1 = q^n - 1$  marks ( $\neq 0$ ) of the field  $F[s = q^n]$  are the  $s - 1$  powers of a mark  $\rho$  belonging to the exponent  $s - 1$ , or say, of a primitive root  $\rho$  of the equation  $X^{s-1} - 1 = 0$ , or of the field  $F[s = q^n]$  itself. The multiplicative-group of the  $s - 1$  marks is cyclic.

(34°) Similarly: Any equation of the form

$$X^f - 1 = 0 ,$$

where  $f$  is any divisor of  $s - 1 = q^n - 1$ , has in the  $F[q^n]$   $f$  roots (32°), of which  $\phi(f)$  are primitive roots.

$\rho$  being a primitive root of the equation

$$X^{s-1} - 1 = 0,$$

then  $\rho^{\frac{s-1}{f}}$  is a primitive root of the equation

$$X^f - 1 = 0.$$

(35°) All the results reached may be applied to any field  $F[q']$ , in particular to any  $F[q']$  included within the  $F[q^n]$ , where then (26°)  $l$  is a divisor of  $n$ . The marks of such a  $F[q']$  are the  $q^l$  roots of the equation (31°, 32°)

$$X^{q^l} - X = X(X^{q^{l-1}} - 1) = 0.$$

Our field  $F[q^n]$  containing one  $F[q']$  contains *no other*  $F[q']$ .

The converse of (35°) is true; see (39°).

(36°) Every mark  $\nu$  ( $\neq 0$ ) of the  $F[q^n]$  is the root of an equation

$$F_k(X) \equiv \sum_{i=1}^k c_i X^i = 0 \quad (c_k=1),$$

belonging to and irreducible in the  $F[q]$  of degree  $k$  equal to the rank of the mark with respect to the  $F[q]$ ; see (22°).

The greatest common divisor of  $X^{q^n} - X$  and  $F_k(X)$  both of which belong to the  $F[q]$  does itself also belong to the  $F[q]$ ; but  $F_k(X)$  is irreducible in the  $F[q]$ ; the divisor is accordingly  $F_k(X)$  itself, or some mark independent of  $X$ . The process necessary to determine this divisor may however be interpreted *also in the*  $F[q^n]$ , in which case  $X - \nu$  is recognized as a common factor (31°, 22°, footnote). Whence in the  $F[q]$   $X^{q^n} - X$  is exactly divisible by every such expression  $F_k(X)$ . Here  $k$  is a divisor of  $n$  (26°, 22°).

(37°) First converse of (36°). Every factor  $F_k(X)$  of degree  $k$  of  $X^{q^n} - X$  which belongs to and is irreducible in the  $F[q]$  has in the  $F[q^n]$   $k$  linear factors (31°, 32°)

$$F_k(X) \equiv c_k \prod_{i=1}^k (X - \nu_i).$$

The mark  $\nu_i$  defines (23°) a Galois-field  $GF[q^k]$  of order  $q^k$  contained in the  $F[q^n]$ . Whence  $k$  is a divisor of  $n$  (22°).

(38°) Second converse of (36°). Every expression  $F_k(X)$  of degree  $k$ , a divisor of  $n$ , belonging to and irreducible in the  $F[q]$  occurs as a factor of  $X^{q^n} - X$ . For every such  $F_k(X)$  serves (by the Galois theory) to define a  $GF[q^k]$ ; whence the  $F_k(X)$  is a factor of  $X^{q^n} - X = X(X^{q^{n-1}} - 1)$  (36°) and thus of

$$X(X^{q^{n-1}} - 1) = X^{q^n} - X.$$

(39°) Galois-fields  $GF[p^k]$  exist (being by the Galois theory defined by such an

$F_*(X), (38^\circ)$  for every prime  $q$  and integer  $k$ . Our  $F[q^n]$  contains one  $(38^\circ, 37^\circ, 23^\circ)$  and only one  $(35^\circ)$   $GF[q^k]$  for every  $k$  a divisor of  $n$ , and no other fields whatever  $(28^\circ, 35^\circ)$ . The marks of such an included  $GF[q^k]$  are the  $q^k$  roots of the equation

$$X^{q^k} - X = 0.$$

$k$  is the rank of the field  $GF[q^k]$   $(21^\circ)$ .

$(40^\circ)$  In particular, for  $k = n$ ,

Every existent field  $F[s]$  is the abstract form of a Galois-field  $GF[q^n]$ ;  $s = q^n$ .

I use always hereafter the notation  $GF[q^k]$  instead of  $F[q^k]$ .

$(41^\circ)$  The marks common to two included fields of ranks  $k_1$  and  $k_2$  respectively constitute a third field, and indeed of rank  $k$  where  $k$  is the greatest common divisor of  $k_1$  and  $k_2$   $(39^\circ, 36^\circ)$ .

$(42^\circ)$  The field of lowest rank containing two included fields of rank  $k_1$  and  $k_2$  respectively is the field of rank  $k'$  where  $k'$  is the least common multiple of  $k_1$  and  $k_2$ .

$(43^\circ)$  The mark  $\nu$  ( $\neq 0$ ) of the  $GF[q^n]$  has the exponent  $e$   $(27^\circ)$ , satisfies the equation  $F_*(X) = 0$   $(22^\circ)$  belonging to and irreducible in the  $GF[q]$ , and defines the  $GF[q^k]$   $(23^\circ, 39^\circ)$ . By the method of proof of  $(36^\circ)$ :  $F_*(X)$  is an exact divisor of  $X^n - 1$  in the  $GF[q]$ ;  $X^n - 1$  is an exact divisor of  $X(X^{q^k-1} - 1) = X^{q^k} - X$  (since  $e$  is a factor of  $q^k - 1$   $(29^\circ)$ );  $X^{q^k} - X$  is an exact divisor of  $X^n - X$  (since the rank  $k$  is a divisor of the rank  $n$   $(26^\circ)$ ). (This accords with  $(36^\circ)$ .)

$(44^\circ)$  The exponent  $e$  of a mark  $\nu$  ( $\neq 0$ ) and the rank  $k$  of the mark  $\nu$  with respect to the  $GF[q]$ , or, what is the same thing, the rank  $k$  of the  $GF[q^k]$  defined by the mark  $\nu$ , are thus related:

$e$  is a divisor of  $q^k - 1$   $(29^\circ)$ , and of  $q^{k'} - 1$  only for  $k'$  a multiple of  $k$ .

$e$  being a divisor of  $q^k - 1$  and of  $q^{k'} - 1$  is a divisor of  $q^l - 1$ , where  $l$  is the greatest common divisor of  $k$  and  $k'$ . Now this  $l$  is  $k$  itself, and so  $k'$  is a multiple of  $k$ . For, since  $e$  is a divisor of  $q^l - 1$ ,  $l \leq k$ ,  $X^e - 1$  is a divisor of  $X^{q^l-1} - 1$  and of  $X^{q^k} - X$ ; the mark  $\nu$  lies in the  $GF[q^l]$   $(12^\circ, 29^\circ, 32^\circ)$ ;  $l$  is a multiple of  $k$   $(23^\circ, 26^\circ)$ , and so in fact equal to  $k$ .

In particular, a primitive root  $(34^\circ)$  of the equation

$$X^{q^k-1} - 1 = 0,$$

where  $k$  is a divisor of  $n$ , is of rank  $k$ . For its  $e$  is  $q^k - 1$ .

$(45^\circ)$  Similarly:

$F_*(X)$  is an exact divisor of  $X^e - 1$ , and of  $X^{e'} - 1$  only for  $e'$  a multiple of  $e$ .

$F_*(X)$  is an exact divisor of  $X^{q^k-1} - 1$ , and of  $X^{q^{k'}-1} - 1$  only for  $k'$  a multiple of  $k$ .

$(46^\circ)$  Similarly, using also the method of  $(36^\circ)$ :

The mark  $\nu$  is a root of the equation  $X^e - 1 = 0$ , and of  $X^{e'} - 1 = 0$  only for  $e$  a multiple of  $e$ .

The mark  $\nu$  is a root of the equation  $X^{q^k-1} - 1 = 0$ , and of  $X^{q^{k'}-1} - 1 = 0$  only for  $k'$  a multiple of  $k$ .



(47°) The equation

$$F_k(X) = 0$$

is in the  $GF[q^k]$  completely reducible, having the roots

$$X = \nu, \nu^q, \nu^{q^2}, \dots, \nu^{q^{k-1}} (\nu^{q^k} = \nu).$$

These  $k$  marks are distinct, for if two were equal, the mark  $\nu$  would satisfy an equation  $X^{q^{k'}} - X = 0$  with  $k' < k$ , contrary to (46°).

That they are roots of the equation depends upon the following lemma:

*Lemma.*—The rational integral function

$$F_k(X) \equiv \sum_{i=0}^k c_i X^i,$$

belongs to the  $GF[q']$  and satisfies the equation

$$\{F_k(X)\}^{q^h} \equiv F_k(X^{q^h}),$$

where  $h$  is any positive integer.

This is proved for the general  $h$ , if proved for  $h = 1$ .

$$\{F_k(X)\}^q \equiv \left\{ \sum_{i=0}^k c_i X^i \right\}^q \equiv \sum_{i=0}^k c_i^q X^{iq} \equiv \sum_{i=0}^k c_i (X^q)^i \equiv F_k(X^q),$$

since the multinomial coefficients for the product terms vanish in the  $GF[q]$  owing to the presence of the factor  $q$ , and since  $c_i^q = c_i$  (Fermat's Theorem; (31°)).

(48°) The mark  $\lambda$  being now *any* mark in the  $GF[q^k]$  it defines a  $GF[q']$ ,  $l$  a certain divisor of  $k$ . Then  $\lambda$  satisfies the equation

$$X^{q^l} - X = 0.$$

*Def.*—The  $k$  marks

$$\lambda' = \lambda^{q^0}, \lambda^{q^1}, \lambda^{q^2}, \dots, \lambda^{q^{k-1}} (\lambda^{q^k} = \lambda)$$

are called *conjugate with respect to the  $GF[q]$* .

These  $k$  conjugate marks are, in view of  $\lambda^{q^l} = \lambda$ , the  $l$  marks

$$\lambda^1 = \lambda^{q^0}, \lambda^{q^1}, \lambda^{q^2}, \dots, \lambda^{q^{l-1}},$$

$k/l$  times repeated. These  $l$  marks are the roots of an equation

$$F_l(X) \equiv \prod_{i=0}^{l-1} (X - \lambda^{q^i}) = 0,$$

belonging to and irreducible in the  $GF[q]$  (22°, 47°).

The expression

$$F_k(X) \equiv \prod_{i=0}^{k-1} (X - \lambda^{q^i})$$

in the  $GF[q]$  has the decomposition

$$F_k(X) \equiv \{F_l(X)\}^{\frac{k}{l}}.$$

It would be interesting, and for other allied investigations necessary, to study in detail the properties of the  $GF[q^n]$  with respect to *any* included  $GF[q']$ , as we have studied the properties with respect to the  $GF[q']$ .

*Concerning squares and not-squares in the  $GF[q^n]$ .*

(49°) Let  $\rho$  be a primitive root of the  $GF[q^n]$ , that is, of the equation (33°)

$$X^{s-1} - 1 \equiv X^{q^n-1} - 1 = 0.$$

(a) If  $q > 2$ , and so  $s-1 = \text{even} = 2t$ , then  $\rho^t = -1 \neq +1$ ; the  $\begin{cases} \text{even} \\ \text{odd} \end{cases}$  powers of  $\rho$  are  $\begin{cases} \text{squares} \\ \text{not-squares} \end{cases}$ ; every square  $\sigma = \rho^{2^k}$  has two square roots  $\rho^k$  and  $\rho^{k+t} = -\rho^k$ . With respect to the multiplication and division of squares and not-squares the well-known laws of the theory of numbers [that is, for the particular case of the  $GF[q]$ ,  $n=1$ ] hold.

(b) If  $q = 2$ , and so  $s-1 = \text{odd} = 2t+1$ , then  $\rho^t$  is also a primitive root of the  $GF[q^n = 2^n]$  (since 2 is prime to  $s-1$ ); every mark is a square, and has only one square root. (We notice that  $-1 = +1$  in this  $GF[2^n]$ .)

Thus in the  $GF[q^n]$

$$\text{If } \begin{cases} q > 2 \\ q = 2 \end{cases} \text{ there are } \begin{cases} \frac{1}{2}(s-1) \text{ squares, } \frac{1}{2}(s-1) \text{ not-squares.} \\ s-1 = 2^n-1 \text{ squares.} \end{cases}$$

(50°) If  $q > 2$ , the not-squares of any  $GF[q']$  are in an including  $GF[q^n]$  not-squares or squares according as the rank  $n/l = h$  of the  $GF[q^n]$  with respect to the  $GF[q']$  is odd or even.

$$\rho_1 = \rho^u \text{ where } u = \frac{q^n-1}{q'-1} \text{ is a primitive root of the } GF[q'] \text{ (34°).}$$

The marks ( $\neq 0$ ) of the  $GF[q']$  are given by  $\rho_1^v = \rho^{uv}$  ( $v = 0, 1, \dots, (l-1)$ ). Let  $\rho_1^v$  be in  $GF[q']$  a not-square;  $v$  is odd;  $\rho_1^v = \rho^{uv}$  is in  $GF[q^n]$  a not-square or a square according as  $uv$ , that is, as  $u$ , is odd or even, that is, in fact, according as  $h = n/l$  is odd or even. For by the equation

$$u = \frac{q^n-1}{q'-1} = \frac{q^{h-1}-1}{q'-1} = \sum_{i=0}^{h-1} q^{i},$$

we exhibit  $u$  as the sum of  $h$  odd terms.

(51°) In the  $GF[q^n]$  the additive-group of lowest rank containing all the squares\* (or not-squares) is the  $GF[q^n]$  itself. This is true for  $q = 2$  (49°). For  $q > 2$ , if this were not true, the  $\frac{1}{2}(q^n-1)$  squares and the 0 would constitute a  $GF[q' = \frac{1}{2}(q^n+1)]$  where  $q' = q$ , *quod* the fundamental additive-period (15°), and this is impossible.

\* BURNSIDE (*loc. cit.*, p. 118) is the first to state this theorem. His proof is inadequate.

§3. THE GALOIS-FIELD OF ORDER  $s = q^{2n}$ 

*Notation.*—The  $GF[q^{2n}]$  contains one  $GF[q^n]$  (§ 2, 39°). Hereafter in this paper I shall use the small Roman letters  $a, b, c, \dots$  to denote integers and also marks of the  $GF[q]$ , the small Greek letters  $\alpha, \beta, \gamma, \dots$  to denote marks of the  $GF[q^n]$ , and the large Roman letters  $A, B, C, \dots$  to denote marks of the  $GF[q^{2n}]$ .

*Def.*—In the  $GF[q^{2n}]$  the two marks  $A, B$  where  $B = A^{q^n}$ ,  $A = B^{q^n}$  (since  $A^{q^{2n}} = A$ ) are called *conjugate with respect to the  $GF[q^n]$* . The notation  $\bar{A}$  for the conjugate of  $A$  is used.

(1°) If and only if  $A$  belongs to the  $GF[q^n]$ , does  $\bar{A} = A$ .

(2°)  $\overline{AC} = \bar{A}\bar{C}$ ,  $\overline{A'} = (\bar{A})'$ ,  $\bar{\bar{A}} = A$ ,  $\overline{\alpha A + \beta B + \gamma C} = \alpha \bar{A} + \beta \bar{B} + \gamma \bar{C}$ .

(3°) Every mark  $A$  of the  $GF[q^{2n}]$  and not of the  $GF[q^n]$  is the root of an equation

$$F_2(X) \equiv X^2 - \kappa X + \lambda = 0,$$

belonging to and irreducible in the  $GF[q^n]$ . The two roots are  $X = A$  and  $X = A^{q^n} = \bar{A}$ .

For:  $A + \bar{A} = K$  and  $A\bar{A} = L$  belong to the  $GF[q^n]$ , since (2°, 1°)

$$\begin{aligned}\bar{K} &= \overline{A + \bar{A}} = \bar{A} + A = K, \\ \bar{L} &= \overline{A\bar{A}} = \bar{A}A = L.\end{aligned}$$

We may then set  $A + \bar{A} = \kappa$ ,  $A\bar{A} = \lambda$  and have in

$$X^2 - \kappa X + \lambda = 0$$

an equation belonging to the  $GF[q^n]$  with the roots  $X = A, \bar{A}$  and irreducible in the  $GF[q^n]$  since its roots do not belong to  $GF[q^n]$ .

(4°) Converse of (3°). Every quadratic equation

$$F_2(X) \equiv X^2 - \kappa X + \lambda = 0$$

belonging to and irreducible in the  $GF[q^n]$  is reducible in the  $GF[q^{2n}]$ , having as roots a pair of conjugate marks  $A, \bar{A}$ .

(a) There are in all  $q^{2n}$  distinct quadratic equations belonging to the  $GF[q^n]$ .

(b) Of these  $\frac{1}{2}(q^{2n} + q^n)$  are reducible in the  $GF[q^n]$ , viz.,  $q^n$  with equal and  $\frac{1}{2}(q^n(q^n - 1))$  with unequal roots.

(c) There are  $\frac{1}{2}(q^{2n} - q^n)$  irreducible ones which are exactly the equations (3°) whose roots are respectively the  $\frac{1}{2}(q^{2n} - q^n)$  pairs of conjugate marks  $A, \bar{A}$ .

Note that every mark  $\beta$  a not-square in the  $GF[q^n]$  is a square in the  $GF[q^{2n}]$  (§ 2, 50°).

(5°) Since  $q^n + 1$  is a divisor of  $q^{2n} - 1$ , the equation

$$X^{q^n+1} - 1 = 0$$

has in the  $GF[q^{2n}]$  as roots the  $q^n + 1$  successive powers of a primitive root  $J$  (§ 2, 34°),

where  $J = R^{q^n-1}$ ,  $R$  being a primitive root of the  $GF[q^n]$ . Any mark  $A$  whose conjugate is its reciprocal

$$A\bar{A} = A^{q^n+1} = 1,$$

is thus a power of  $J$  and with its conjugate satisfies (3°) a quadratic equation of the form,

$$X^2 - \kappa X + 1 = 0.$$

The powers of  $J$  which lie in the  $GF[q^n]$  satisfy likewise the equation

$$X^{q^n-1} - 1 = 0,$$

and so the equation

$$X^2 - 1 = 0;$$

they are then  $J^0 = +1$ , and (for  $q > 2$ ,  $q^n+1$  even)  $J^{\frac{1}{2}(q^n+1)} = -1$ .

Thus there are  $\left\{ \begin{array}{l} \text{if } q > 2, \frac{1}{2}(q^n-1) \\ \text{if } q = 2, \frac{1}{2}q^n = 2^{n-1} \end{array} \right\}$  quadratic equations of the form

$$X^2 - \kappa X + 1 = 0,$$

which belong to and are irreducible in the  $GF[q^n]$ .

I write the one satisfied by  $J$  and  $\bar{J} = J^{-1}$  thus

$$(X - J)(X - \bar{J}) \equiv X^2 - \theta X + 1 = 0, \quad (J; \theta)$$

thereby defining the mark  $\theta$  of the  $GF[q^n]$

$$\theta = J + \bar{J} = J + J^{-1}.$$

(6°) In the  $GF[q^n]$  the mark  $J$  having the period  $e = q^n+1$  defines (§ 2, 23° 39°) the  $GF[q^n]$  itself (§ 2, 44°).

(7°) In the  $GF[q^n]$  the mark  $\theta$  defines the  $GF[q^n]$  itself. For let  $\theta$  define the  $GF[q^l]$  ( $l$  a divisor of  $n$ ). The  $GF[q^n]$  contains a  $GF[q^l]$ , in which the equation belonging to the  $GF[q^l]$

$$X^2 - \theta X + 1 = 0, \quad (J; \theta)$$

is reducible (4°). Hence its root  $J$  belongs to the  $GF[q^l]$ . Thus indeed  $l = n$  (6°).

(8°) There are  $\phi(q^n+1)$  primitive roots  $J$  and thus  $\frac{1}{2}\phi(q^n+1)$  marks  $\theta$ . We do not now need to inquire further into the properties of this system of marks  $\theta$ .

(9°) Any mark  $A$  of the  $GF[q^n]$  may be written in one and only one way in the form

$$A = \gamma + \delta J.$$

(§ 2, 24°). Its conjugate is (2°, 5°)

$$\bar{A} = \gamma + \delta \bar{J}.$$

#### § 4. DEFINITION OF THE GROUP $G_{M(q^n)}^{q^n+1}$

The abstract group  $G_{M(q^n)}$  will be studied under three concrete substitution-group forms  $G_{M(q^n)}^{q^n+1}$ ,  $G'_{M(q^n)}^{q^{2n}+1}$ , and  $G''_{M(q^n)}^{q^{2n}+1}$ . Indeed it will be now *defined* by its concrete form  $G_{M(q^n)}^{q^n+1}$ .

We work in the  $GF[q^n]$  with the mark  $\infty$  adjoined. The mark  $\infty$  shall have with respect to the  $GF[q^n]$  operational properties strictly analogous to those of the algebraic symbol  $\infty$  with respect to the totality of finite algebraic symbols.

Let the variable marks  $\omega, \omega'$  run through this series of  $q^n + 1$  marks. Let  $\alpha, \beta, \gamma, \delta$  be any marks of the  $GF[q^n]$  for which the *determinant*  $\Delta = \alpha\delta - \beta\gamma \neq 0$ . Then the linear fractional substitution

$$V: \quad \omega' = \frac{\alpha\omega + \beta}{\gamma\omega + \delta} \quad (\alpha\delta - \beta\gamma = \Delta \neq 0) \quad (1)$$

makes correspond to every mark  $\omega$  a certain mark  $\omega'$ , and indeed so that  $\omega'$  takes every one of its  $q^n + 1$  values, the *reciprocal substitution* being

$$V^{-1}: \quad \omega = \frac{\frac{\delta}{\Delta}\omega' - \frac{\beta}{\Delta}}{-\frac{\gamma}{\Delta}\omega' + \frac{\alpha}{\Delta}}. \quad (2)$$

$V$  effects then a definite permutation of the  $q^n + 1$  marks  $\omega$  and will be thought of as a notation for a substitution-operation on those  $q^n + 1$  marks.

The notations

$$\left(\frac{\alpha, \beta}{\gamma, \delta}\right), \quad (\alpha, \beta; \gamma, \delta) \quad (3)$$

are used for the substitution  $V$ .

Corresponding to the formulæ

$$\begin{aligned} (V') \quad \omega' &= \frac{\alpha'\omega' + \beta'}{\gamma'\omega' + \delta'}, & (V') \quad \omega' &= \frac{\alpha'\omega + \beta'}{\gamma'\omega + \delta'}, \\ (V) \quad \omega' &= \frac{(\alpha'\alpha' + \beta'\gamma')\omega + (\alpha'\beta' + \beta'\delta')}{(\gamma'\alpha' + \delta'\gamma')\omega + (\gamma'\beta' + \delta'\delta')}, \end{aligned} \quad (4)$$

we have the formula of composition or multiplication of substitutions

$$V'V' = V;$$

$$(\alpha'', \beta''; \gamma'', \delta'')(\alpha', \beta'; \gamma', \delta') = (\alpha''\alpha' + \beta''\gamma', \alpha''\beta' + \beta''\delta'; \gamma''\alpha' + \delta''\gamma', \gamma''\beta' + \delta''\delta'). \quad (5)$$

The determinant of the product of two substitutions is the product of the determinants of the substitutions.

The totality of all such distinct substitutions of determinant  $+1$  constitutes the substitution-group  $G_{M(s)}^{s+1} = G_{M(q^n)}^{q^n+1}$  of order  $M(s) = M(q^n)$  on the  $s + 1$  marks  $\omega$ .

The two substitutions

$$V = (\alpha, \beta; \gamma, \delta), \quad V_\mu = (\mu\alpha, \mu\beta; \mu\gamma, \mu\delta) \quad \left(\alpha\delta - \beta\gamma = \Delta \neq 0, \mu \neq 0\right) \quad (6)$$

are the same in fact. Their determinants are  $\Delta = \alpha\delta - \beta\gamma$ ,  $\Delta_\mu = \mu^2\Delta$ . (a) Case  $q > 2$ . The marks ( $\neq 0$ ) are half squares and half not-squares (§ 2, 49°); and  $\Delta^{-1}$  or  $\Delta^{-1}\rho$  is a square, according as  $\Delta$  is or is not a square ( $\rho$  being a primitive root of the  $GF[q^n]$ ). By choosing then  $\mu = \pm \sqrt{\Delta^{-1}}$  or  $\pm \sqrt{\Delta^{-1}\rho}$  according as  $\Delta$  is or is not a square,  $V_\mu$  has its determinant  $\Delta_\mu = 1$  or  $\rho$  respectively. (b) Case  $q = 2$ . Every

mark ( $\neq 0$ ) is a square (§ 2, 49°). By choosing then  $\mu = \sqrt{\Delta^{-1}}$  (uniquely defined in the  $GF[q^n = 2^n]$ ),  $V_\mu$  has its determinant  $\Delta_\mu = 1$ .

These forms of the substitution for which the determinant is 1 or  $\rho$  ( $q > 2$ ); 1 ( $q = 2$ ) are called *normal* forms of the substitutions.

Every substitution has exactly, if  $q > 2$ , two normal forms, or, if  $q = 2$ , one normal form:

$$\begin{aligned} (q > 2) \quad (\alpha, \beta; \gamma, \delta) &= (-\alpha, -\beta; -\gamma, -\delta) & (\alpha\delta - \beta\gamma = 1 \text{ or } \rho) \\ (q = 2) \quad (\alpha, \beta; \gamma, \delta) & & (\alpha\delta - \beta\gamma = 1) \end{aligned} \quad (7)$$

It is easily proved that the equation of the  $GF[s = q^n]$ ,  $\alpha\delta - \beta\gamma = \epsilon$ , in which  $\epsilon$  is a fixed mark  $\neq 0$ , has  $(s-1)s(s+1)$  solutions  $(\alpha, \beta, \gamma, \delta)$ .

Thus: ( $q > 2$ ). There are  $\frac{1}{2}s(s^2-1)$  substitutions  $(\alpha, \beta; \gamma, \delta)$  having  $\Delta = \alpha\delta - \beta\gamma = \text{square}$  (in normal form,  $\Delta = 1$ ) and  $\frac{1}{2}s(s^2-1)$  having  $\Delta = \text{not-square}$  (in normal form,  $\Delta = \rho$ ).

( $q = 2$ .) There are  $s(s^2-1)$  substitutions  $(\alpha, \beta; \gamma, \delta)$  (in normal form  $\Delta = 1$ ).

Thus clearly the order  $M(s)$ ,  $s = q^n$ , of the group  $G_{M(s)}^{s+1}$  (embracing the totality of substitutions with determinant 1) is\*

$$\frac{s(s^2-1)}{2; 1} = \frac{q^n(q^{2n}-1)}{2; 1} = M(s) = M(q^n) \quad (q > 2; q = 2). \quad (8)$$

The group of all the substitutions with determinants not 0 has the order  $(2; 1) M(s)$ , according as ( $q > 2$ ;  $q = 2$ ). For  $q > 2$  the  $G_{M(s)}^{s+1}$  contains the  $G_{M(s)}$  as a self-conjugate subgroup of index 2. Since every mark  $\Delta$  ( $\Delta \neq 0$ ) of the  $GF[q^n]$  is a square in the  $GF[q^{2n}]$  (§ 2, 50°), the  $G_{M(s)}$  ( $q > 2$ ) is a subgroup of the  $G_{M(s^2=q^{2n})}$  of the  $GF[q^{2n}]$ . Hence, we easily extend (section by section) the investigation of the subgroups of the  $G_{M(s)}$  to the subgroups of the  $G_{(2; 1) M(s)}$ .

#### § 5. DEFINITION OF THE GROUPS $G_{M(q^n)}^{q^{2n}+1}$ , $G_{M(q^n)}^{q^{2n}+1}$ , TWO "IMAGINARY" FORMS OF THE $G_{M(q^n)}$

We work in the  $GF[q^{2n}]$  with a mark  $\infty$  adjoined, and let the variable marks  $W$ ,  $W'$  run through this series of  $q^{2n} + 1 = s^2 + 1$  marks. ( $s$  is used hereafter uniformly for  $q^n$ .)

We consider the group  $G_{(2; 1) M(s)}^{s^2+1}$  ( $q > 2$ ;  $q = 2$ ) of all substitutions of the form

$$\left( W' = \frac{A W + B}{C W + D} \right) \text{ or } \left( \frac{A, B}{C, D} \right) \text{ or } (A, B; C, D) \quad (AD - BC \neq 0). \quad (1)$$

This group contains the subgroup  $G_{(2; 1) M(s)}^{s^2+1}$  of all substitutions of the form

$$\left( W' = \frac{\alpha W + \beta}{\gamma W + \delta} \right) \text{ or } \left( \frac{\alpha, \beta}{\gamma, \delta} \right) \text{ or } (\alpha, \beta; \gamma, \delta) \quad (\Delta = \alpha\delta - \beta\gamma \neq 0), \quad (2)$$

\* We shall need to discriminate frequently between the cases  $q > 2$ ,  $q = 2$ . The compact notation used for such discrimination should cause no confusion.

which is holoelectrically isomorphic with the substitution-group  $G_{(s;1)M(s)}^{s+1}$  (§4) on the  $s+1$  marks  $\omega$ ; indeed, it is an intransitive group on the  $s^2+1$  marks  $W$ , the  $s+1$  marks  $\omega$  forming one system of intransitivity.

The group  $G_{(s;1)M(s)}^{s+1}$  contains also the substitution

$$R = (J, 1; \bar{J}, 1) \quad (3)$$

of determinant  $J - \bar{J} \neq 0$  (§3, 1°, 5°). We have (§3, 5°)

$$J^2 = \theta J - 1; \bar{J}^2 = \theta \bar{J} - 1. \quad (4)$$

$$J\bar{J} = 1; J + \bar{J} = \theta; (J - \bar{J})^2 = \theta^2 - 4 \neq 0. \quad (5)$$

$$R^{-1} = ((J - \bar{J})^{-1}, -(J - \bar{J})^{-1}; -\bar{J}(J - \bar{J})^{-1}, J(J - \bar{J})^{-1}). \quad (6)$$

$R$  transforms<sup>10</sup> the group  $G_{(s;1)M(s)}^{s+1} \{V\}$  into an holoelectrically isomorphic group  $G'_{(s;1)M(s)}^{s+1} \{V'\} = R\{V\}R^{-1} = \{V' = RVR^{-1}\}$ , where

$$V = (\alpha, \beta; \gamma, \delta) \quad (\Delta = \alpha\delta - \beta\gamma \neq 0); \quad (7)$$

$$V' = \left( \frac{A, B}{B, A} \right) = \left( \frac{\kappa + \lambda J, \mu + \nu J}{\mu + \nu \bar{J}, \kappa + \lambda \bar{J}} \right) \quad (\Delta' = A\bar{A} - B\bar{B} = \Delta), \quad (8)$$

$$\kappa = \delta + \frac{(\beta - \gamma)\theta - 2(\alpha - \delta)}{\theta^2 - 4}, \quad (9)$$

$$\lambda = + \frac{(\alpha - \delta)\theta - 2(\beta - \gamma)}{\theta^2 - 4},$$

$$\mu = - \frac{(\beta - \gamma)\theta - 2(\alpha - \delta)}{\theta^2 - 4},$$

$$\nu = \beta - \frac{(\alpha - \delta)\theta - 2(\beta - \gamma)}{\theta^2 - 4};$$

$$\Delta' = A\bar{A} - B\bar{B} = \kappa^2 + \lambda^2 - \mu^2 - \nu^2 + (\kappa\lambda - \mu\nu)\theta = \alpha\delta - \beta\gamma = \Delta. \quad (10)$$

This group  $G'_{(s;1)M(s)}^{s+1}$  includes *every* substitution of the form  $V'$  (8). The  $V'$  is determined from the  $V$  by the formulæ

$$\alpha = +\kappa - \mu + \lambda\theta, \quad (11)$$

$$\beta = +\lambda + \nu,$$

$$\gamma = -\lambda + \nu + \mu\theta,$$

$$\delta = +\kappa + \mu;$$

$$\Delta = \alpha\delta - \beta\gamma = \kappa^2 + \lambda^2 - \mu^2 - \nu^2 + (\kappa\lambda - \mu\nu)\theta = A\bar{A} - B\bar{B} = \Delta'. \quad (12)$$

The group  $G'_{(s;1)M(s)}^{s+1} \{V'\}$  (8) ( $\Delta' = 1$ ) is the first "imaginary" form of the  $G_{M(s)}$ . For  $n=1$ ,  $q>2$  this form was introduced (by Serret, Mathieu, and then) by Gierster, *loc. cit.*, p. 327, and indeed, as we have done for the general case, by introducing the broader group  $G_{(q)M(q)}^{q+1}$  within which  $G_{M(q)}^{q+1}$  ( $\Delta=1$ ) and the  $G'_{M(q)}^{q+1}$  ( $\Delta'=1$ ) are conjugate

<sup>10</sup>  $R, V$  being two operators of a group,  $R$  transforms  $V$  into the conjugate operator  $V \rightarrow RVR^{-1}$ . This definition is preferable to the other,  $V \rightarrow R^{-1}VR$ , when as here in the

$G_{M(s)}^{s+1}$  the operators are substitutions on the  $s+1$  marks operating from the right to the left.

subgroups (*loc. cit.*, p. 331). Gierster (*loc. cit.*, p. 328) and Klein-Fricke (*loc. cit.*, p. 425) use transforming substitutions which depend upon the square root in the  $GF[q^2]$  of a not-square in the  $GF[q^1]$ . Since, however, not-squares do not exist in  $GF[s=q^n]$  with  $q=2$ , we have made use of a transforming substitution  $R$  of a different type which does exist for the general  $n$  with  $q=2$  as well as with  $q>2$ .

The group  $G_{(\alpha;1)M(q^n)}^{s^{s+1}}$  contains the substitution

$$R' = (K, 0; 0, 1), \quad (13)$$

where  $K$  is a particular square root of  $J$ , so that

$$K\bar{K} = -1, \quad R'^{-1} = (-\bar{K}, 0; 0, 1). \quad (14)$$

Under  $R'$  the group  $G_{(\alpha;1)M(q)}^{s^{s+1}} \{V'\}$  (8) transforms into the holodrically isomorphic group  $G_{(\alpha;1)M(q)}^{s^{s+1}} \{V''\}$  of all substitutions of the form

$$V'' = (A, B; -\bar{B}, \bar{A}) \quad (\Delta' = A\bar{A} + B\bar{B} = \Delta). \quad (15)$$

The group  $G_{M(q)}^{s^{s+1}} \{V''\}$  ( $\Delta' = 1$ ) is the (new) second imaginary form<sup>11</sup> of the  $G_{M(q)}$ . This form is found more convenient than the first form (§§ 6, III, 15, 16), since it contains  $T = (0, -1; 1, 0)$ .

We notice the three corresponding substitutions:

$$V_\theta = (\theta, 1; -1, 0), \quad V'_\theta = V''_\theta = D'_\theta = (J, 0; 0, \bar{J}), \quad (16)$$

where in general we write

$$D_\Delta = (A, 0; 0, 1). \quad (16')$$

Within the  $G_{(\alpha;1)M(q^n)}^{s^{s+1}}$  the three conjugate groups,

$$(a) G_{(\alpha;1)M(q)}^{s^{s+1}} (2), \quad (b) G_{(\alpha;1)M(q)}^{s^{s+1}} (8), \quad (c) G_{(\alpha;1)M(q)}^{s^{s+1}} (15)$$

contain respectively all substitutions of non-vanishing determinant of the form

$$V = (\alpha, \beta; \gamma, \delta), \quad V' = (A, B; \bar{B}, \bar{A}), \quad V'' = (A, B; \bar{B}, \bar{A}). \quad (17)$$

The necessary and sufficient conditions that a substitution

$$U = (A, B; C, D) = (AM, BM; CM, DM) \quad (AD - BC \neq 0; M \neq 0)$$

belong to the respective groups are:

$$\begin{aligned} (a) \quad & AB^* = A^*B, \quad AC^* = A^*C, \quad AD^* = A^*D, \\ & CD^* = C^*D, \quad DB^* = D^*B, \quad BC^* = B^*C; \\ (b) \quad & A^{s+1} = D^{s+1}, \quad B^{s+1} = C^{s+1}, \quad AC^* = D^*B; \\ (c) \quad & A^{s+1} = D^{s+1}, \quad B^{s+1} = C^{s+1}, \quad AC^* = -D^*B. \end{aligned} \quad (18)$$

Thus, for instance, the substitution

$$(1, 1; 1, -1) = (\sqrt{\rho}, \sqrt{\rho}; \sqrt{\rho}, -\sqrt{\rho})$$

of determinant  $-2$ ,  $-2\rho$  belongs (for  $q>2$ ) both to the real and to the second imaginary  $G_{2M(q)}$ , while it belongs to the real or to the second imaginary  $G_{M(q)}$  according as  $-2$  is a square or a not-square in the  $GF[s]$ .

<sup>11</sup>In the recent automorphic-function literature a similar type of substitution is prominent. For instance, FRICKE, "Zur gruppentheoretische Grundlegung der auto-

morphen Functionen," *Mathematische Annalen*, Vol. XLII (1893), p. 564. This second imaginary form of the group was independently discovered by Dickson during the year 1899.



### §6. THE INDIVIDUAL OPERATORS AND THE LARGEST CYCLIC AND COMMUTATIVE SUBGROUPS OF THE $G_{M(s)}$

To summarize at once the conclusions of this section:

Of the  $G_{M(s)}$  every operator (the identity excepted) determines and lies in one and only one largest<sup>12</sup> commutative subgroup. These subgroups constitute three different sets, the groups of each set being conjugate with one another under the  $G_{M(s)}$ .

(I)  $s + 1$  conjugate commutative  $G_{s+1}$ . These  $s^2 - 1$  operators are all of period  $q$ . For  $q = 2$  these are all conjugate operators. For  $q > 2$  they separate into two sets of  $\frac{1}{2}(s^2 - 1)$  conjugate operators. The  $q - 1$  operators of a cyclic group  $G_q$  ( $q > 2$ ) belong  $\left\{ \begin{array}{l} n \text{ odd, half to one and half to the other} \\ n \text{ even, all to the same} \end{array} \right\}$  set of conjugate operators.

(II)<sup>12</sup>  $\frac{1}{2}s(s + 1)$  conjugate cyclic groups  $G_{\frac{s+1}{2}}(q > 2; q = 2)$ .

(III)<sup>12</sup>  $\frac{1}{2}s(s - 1)$  conjugate cyclic groups  $G_{\frac{s-1}{2}}(q > 2; q = 2)$ .

(IV)<sup>12</sup> Two substitutions  $(A, B; C, D), (A', B'; C', D')$ , of determinant 1 of the  $G_{M(s)}$  in its real or first imaginary or second imaginary form (§5) (neither substitution being the identity) are conjugate if and only if  $(A + D)^2 = (A' + D')^2$ . But—the substitutions of set I all have  $(A + D)^2 = 4$  and for  $q > 2$  separate into two sets of conjugate substitutions.

(V)<sup>12</sup> For  $q > 2$ , for the substitutions of set

$$\left. \begin{array}{l} \text{(I)}^{12} \\ \text{(II)} \\ \text{(III)} \end{array} \right\} (A + D)^2 - 4 \text{ is } \left\{ \begin{array}{l} 0. \\ \text{a square in the } GF[s]. \\ \text{a not-square in the } GF[s]. \end{array} \right.$$

[For the cases  $s = 9; 7$  (the  $G_{M(s)}$  being the  $6!$ -symmetric group  $G_{360}$ ; the simple group  $G_{180}$ ) the cyclic groups (II); (III) of order 4 exist as stated, but every element of order 2 of such a cyclic group lies in and determines not merely this largest commutative subgroup of order 4 but two other four-groups, the three groups of order 4 lying in the dihedron group of order 8 whose cyclic base is the cyclic group in question. These two are the only cases in which an operator determines two or more "largest" commutative subgroups. This follows from the developments of the text in view of the fact that a non-commutative dihedron group of order  $2d$  contains two commutative subgroups of order  $d$  only if  $d = 4$ . The classification of the four-groups is made in §9.]

The operators and groups of the three types will be discussed by considering the  $G_{M(s)}$  in the concrete substitution-group form  $G_{M(s)}^{s+1}$  for types I, II, and  $G_{M(s)}^{s^2+1}$  for type III.

<sup>12</sup>For the cases  $s=q=5; 3$  (the  $G_{M(s)}$  being the icosahedron  $G_{60}$ ; the tetrahedron  $G_{12}$ ) the groups (II); (III) above specified, viz., 15 conjugate  $G_3$  under the  $G_{60}$ ; 3 conjugate  $G_3$  under the  $G_{12}$ , exist, but every one is contained in a larger commutative four-group, these constituting a

system of (II) five conjugate commutative four-groups; (III) one commutative four-group.

<sup>13</sup>Cf. GIERSTER (*loc. cit.*, pp. 326, 327) and BURNSIDE (*loc. cit.*, pp. 122-4).

<sup>14</sup>The characterization V (I) holds also for  $q = 2$ .

I. *The operators of period  $q$ .* — Denote by  $S_\beta$  the substitution of the  $G_{M(\omega)}^{s+1}$ ,

$$S_\beta = (1, \beta; 0, 1) . \quad (1)$$

We have

$$S_\beta S_{\beta_1} = S_{\beta_1 + \beta_2} = S_{\beta_1} S_{\beta_2} ; S_\beta^a = S_{a\beta} \quad (2)$$

$$S_\beta^q = S_{q\beta} = S_0 = I ; S_\beta^{-1} = S_{-\beta} = S_\beta^{q-1} ; \quad (3)$$

where  $I$  denotes the identical substitution,

$$I = (1, 0; 0, 1) . \quad (4)$$

The totality of  $s = q^n$  substitutions  $S_\beta$  constitutes a commutative group  $G_s^{(\omega)} \{S_\beta\}$  of order  $s = q^n$ ; it contains all and only the substitutions leaving the single mark  $\omega = \infty$  invariant. Every substitution (except the identity) is of period  $q$ . There are  $(s-1)/(q-1)$  cyclic  $G_q$  in the  $G_s^{(\omega)}$ .

To study the conjugacy of these substitutions and groups under the  $G_{M(\omega)}^{s+1}$  we transform  $S_\mu (\mu \neq 0)$  by  $V = (\alpha, \beta; \gamma, \delta)$ , and obtain

$$VS_\mu V^{-1} = \left( \frac{1 - \alpha\gamma\mu, \alpha^2\mu}{-\gamma^2\mu, 1 + \alpha\gamma\mu} \right) . \quad (5)$$

This transformed substitution belongs to the  $G_s^{(\omega)}$  if and only if in  $V\gamma = 0$ , and in fact  $V = (\alpha, \beta; 0, \alpha^{-1})$ , does transform  $S_\mu$  into  $S_{\alpha^2\mu}$ , while in particular any  $S_\beta = (1, \beta; 0, 1)$  transforms  $S_\mu$  into itself.

Within the  $G_{M(\omega)}^{s+1}$  the substitution  $S_\mu \{\mu \neq 0\}$  is self-conjugate in exactly the  $G_s^{(\omega)} \{S_\beta\}$ , while the  $G_s^{(\omega)}$  is self-conjugate in exactly the  $G_{\frac{s(s-1)}{2; 1}}^{(\omega)} \{(\alpha, \beta; 0, \alpha^{-1})\}$  which contains all the substitutions leaving  $\omega = \infty$  fixed. The order of this last group is  $\frac{s(s-1)}{2; 1}$ ,

for  $\alpha$  has in all  $s-1$  values ( $\neq 0$ ), and  $\beta$  has independently  $s$  values; every substitution is however counted twice if  $q > 2$ . The substitution  $S_\mu$  is conjugate under the  $G_{M(\omega)}^{s+1}$  with the substitutions  $S_{\alpha^2\mu}$ , i. e., if  $(q > 2; q = 2)$ , with (only half; all) the  $s-1$  substitutions of the  $G_s^{(\omega)}$  (§ 2, 49°). If  $q > 2$ , the  $s-1$  substitutions of the  $G_s^{(\omega)}$  separate into two sets of  $\frac{1}{2}(s-1)$  conjugate substitutions; under the  $G_{M(\omega)}$  they are however all conjugate, since by  $D_p$  (§ 4 (9))  $S_1$  transforms to  $S_p$ . The  $q-1$  substitutions of a cyclic group  $G_q \{S_{\alpha\mu}\}$  ( $\alpha \neq 0$ ) belong half to one and half to the other set, if  $n$  is odd, since then only half the integral marks  $\alpha$  are in the  $GF[q^n]$  square (§ 2, 50°), while if  $n$  is even, they all belong to the same set, since all the marks  $\alpha$  are in the  $GF[q^n]$  squares.

In the  $G_{M(\omega)}^{s+1}$  there are

$$\left\{ M(s) = \frac{s(s^2-1)}{2; 1} \right\} \div \frac{s(s-1)}{2; 1} = s+1$$

conjugate commutative groups  $G_s^{(\omega)}$ . Each of these is defined by any substitution lying in it (the identity excepted) as the group in which that substitution is self-conjugate. The  $s+1$  groups have the identity in common, but otherwise have quite distinct substitutions all of period  $q$ ,  $s^2-1$  in all.

$G_i^{(\kappa)}$  is self-conjugate in a group  $G_{\frac{s(s-1)}{2}; 1}^{(\kappa)}$ . These  $s+1$  groups  $G_{\frac{s(s-1)}{2}; 1}^{(\kappa)}$  are under the  $G_{M(\omega)}$  a set of conjugate groups. They are distinct, since  $G_{\frac{s(s-1)}{2}; 1}^{(\omega)}$  contains substitutions  $S_\beta$  lying in no other  $G_{\frac{s(s-1)}{2}; 1}^{(\kappa)}$ .

II. *The operators of periods divisors of  $\frac{s-1}{2}$ .*—Denote by  $P$  the substitution of the  $G_{M(\omega)}^{s+1}$

$$P = (\rho, 0; 0, \rho^{-1}), \quad (8)$$

where  $\rho$  is a primitive root of the field  $GF[s]$  (§2, 33°). We have as powers of  $P$

$$P^\sigma = (\rho^\sigma, 0; 0, \rho^{-\sigma}) \quad (\sigma = 0, \pm 1, \pm 2, \dots) \quad (7)$$

$\rho$  belongs to the exponent  $s-1$ . The cyclic group generated by  $P$  has order  $\frac{s-1}{2}$  ( $q > 2$ ;  $q = 2$ ); we denote it by  $G_{\frac{s-1}{2}; 1}^{(\omega)} \{P^\sigma\}$ , since it contains all and only the substitutions leaving  $\omega = \infty$  and  $\omega = 0$  each fixed. The substitutions  $(\alpha, \beta; \gamma, \delta)$ ,  $(\alpha\delta - \beta\gamma = 1)$  of this group are defined by the equations  $\beta = 0, \gamma = 0$ . We write also

$$P_\alpha = (\alpha, 0; 0, \alpha^{-1}) = P_{-\alpha} \quad (\alpha \neq 0). \quad (7')$$

This group  $G_{\frac{s-1}{2}; 1}^{(\omega)}$  contains a substitution of period two if and only if  $s = q^n$  has the form  $s = q^n = 4h+1$ ; in this case

$$P^h = P^{\frac{s-1}{2}} = (\sqrt{-1}, 0; 0, -\sqrt{-1}) \quad (8)$$

exists and has the period two.

Any substitution  $V$  of the  $G_{M(\omega)}^{s+1}$

$$V = (\alpha, \beta; \gamma, \delta) \quad (\alpha\delta - \beta\gamma = 1) \quad (9)$$

transforms  $P^\sigma \neq I$  ( $\rho^\sigma \neq \pm 1$ ; and so  $\rho^\sigma - \rho^{-\sigma} \neq 0$ ) into

$$V P^\sigma V^{-1} = \left( \frac{\alpha\delta\rho^\sigma - \beta\gamma\rho^{-\sigma}}{\gamma\delta(\rho^\sigma - \rho^{-\sigma})}, \frac{-\alpha\beta(\rho^\sigma - \rho^{-\sigma})}{\alpha\delta\rho^{-\sigma} - \beta\gamma\rho^\sigma} \right). \quad (10)$$

This transformed substitution belongs to the cyclic group  $G_{\frac{s-1}{2}; 1}^{(\omega)} \{P^\sigma\}$  if in  $V$   $\alpha\beta = 0$  and  $\gamma\delta = 0$ . We have also  $\alpha\delta - \beta\gamma = 1$ . Two cases arise.

First case:  $\beta = \gamma = 0$ .  $V = (\alpha, 0; 0, \alpha^{-1})$  itself belongs to the cyclic group  $G_{\frac{s-1}{2}; 1}^{(\omega)}$  and of course transforms every  $P^\sigma$  of the group into itself.

Second case:  $\alpha = \delta = 0$ .  $V = (0, \beta; -\beta^{-1}, 0)$  transforms  $P^\sigma$  into  $P^{-\sigma}$ , which is distinct from  $P^\sigma$  unless  $P^\sigma$  is of period two.

Within the  $G_{M(\omega)}$  the cyclic  $G_{\frac{s-1}{2}; 1}^{(\omega)}$  is self-conjugate in exactly the dihedron group  $G_{\frac{s-1}{2}; 1}^{(\omega)}$  composed of the totality of substitutions having the forms  $(\alpha, 0; 0, \alpha^{-1})$ ,  $(0, \beta; -\beta^{-1}, 0)$ .

Within the  $G_{M(s)}$  a substitution  $P^s$  is self-conjugate in exactly the  $G_{\frac{s-1}{2}; 1}^{(\infty, 0)}$ , except in the case  $s = q^n = 4h + 1$ , in which  $P^{\frac{s-1}{2}}$  is self-conjugate in exactly the dihedron  $G_{\frac{s-1}{2}}^{(\infty, 0)}$ ; this dihedron group is commutative only if it is a four-group ( $s = q^1 = 5$ ).

In the  $G_{M(s)}$  there are  $\left\{ M(s) = \frac{s(s^2-1)}{2; 1} \right\} \div 2 \frac{s-1}{2; 1} = \frac{1}{2}s(s+1)$  conjugate cyclic groups  $G_{\frac{s-1}{2}; 1}^{(\kappa\lambda)} = G_{\frac{s-1}{2}; 1}^{(\lambda\kappa)}$  ( $\kappa \neq \lambda$ ). Each of these is *defined* by any substitution lying in it (the identity excepted) as *the largest* cyclic group containing that substitution. The  $\frac{1}{2}s(s+1)$  groups have the identity in common, but otherwise have quite distinct substitutions, of periods divisors of  $\frac{s-1}{2; 1}$ , in all

$$\frac{s \cdot s + 1}{2} \left( \frac{s-1}{2; 1} - 1 \right) = \frac{s(s+1)(s-3; s-2)}{4; 2}$$

substitutions.

In the  $G_{M(s)}$  there are  $M(s) \div \frac{s-1}{2; 1} = s(s+1)$  substitutions conjugate to  $P^s$  of which two lie in every  $G_{\frac{s-1}{2}; 1}^{(\kappa\lambda)}$  (for instance, in  $G_{\frac{s-1}{2}; 1}^{(\infty, 0)}$  lie  $P^s$  and  $P^{-s}$ ). However, for  $s = q^n = 4h + 1$ , there are only  $\frac{1}{2}s(s+1)$  substitutions conjugate to  $P^{\frac{s-1}{2}}$ , one lying in every  $G_{\frac{s-1}{2}}^{(\kappa\lambda)}$ .

$G_{\frac{s-1}{2}; 1}^{(\kappa)} \text{ and } G_{\frac{s-1}{2}; 1}^{(\lambda)}$  have  $G_{\frac{s-1}{2}; 1}^{(\kappa\lambda)}$  in common.  $G_{\frac{s-1}{2}; 1}^{(\infty)}$  is made up of  $G_{\frac{s-1}{2}; 1}^{(\infty)}$  and the  $s$   $G_{\frac{s-1}{2}; 1}^{(\infty, \kappa)}$  ( $\kappa$  any  $\neq 0$ );  $S_s$  transforms  $G_{\frac{s-1}{2}; 1}^{(\infty, \kappa)}$  to  $G_{\frac{s-1}{2}; 1}^{(\infty, \kappa+\beta)}$ .

III. *The operators of periods divisors of  $\frac{s+1}{2; 1}$ .*—Denote by  $Q$  the substitution of the  $G_{M(s)}^{s+1}$  (§ 5 (16),  $\Delta' = 1$ )

$$(Q = J, 0; 0, \bar{J}), \quad (11)$$

where  $J$  is a primitive root of the equation  $X^{s+1} = 1$  and  $\bar{J} = J^{-1}$  (§ 3, 5°). We have as powers of  $Q$

$$Q^p = (J^p, 0; 0, \bar{J}^p) \quad (p=0, \pm 1, \pm 2, \dots).$$

$J$  belongs to the exponent  $s+1$ . The cyclic group generated by  $Q$  has order  $\frac{s+1}{2; 1}$  ( $q > 2$ ;  $q = 2$ ); denote it by  $G_{\frac{s+1}{2}; 1}^* \{Q^p\}$ . The substitutions  $(A, B; -\bar{B}, \bar{A})$  ( $A\bar{A} + B\bar{B} = 1$ ) of this group are defined by the equation  $B = 0 = \bar{B}$ ; see (§ 3, 5°).

This group  $G_{\frac{s+1}{2}; 1}^*$  contains a substitution of period two if and only if  $s = q^n$  has the form  $s = q^n = 4k - 1$ ; in this case

$$Q^k = Q^{\frac{s+1}{4}} = (\sqrt{-1}, 0; 0, -\sqrt{-1}) \quad (12)$$

exists and has the period two.

We notice particularly that in the  $G_{M(n)}$  operations of period two are always present, being all of the type I, II, or III, according as  $q$  is even ( $=2$ ), or ( $q$  odd)  $q^n$  of the form  $q^n = 4h + 1$  or  $q^n = 4k - 1$ .

Any substitution  $V$  of the  $G_{\frac{s+1}{2}, 1}^*$

$$V = (A, B; -\bar{B}, \bar{A}) \quad (A\bar{A} + B\bar{B} = 1) \quad (13)$$

transforms  $Q^s \neq I$  ( $J^s \neq \pm 1$ ; and so  $J^s - \bar{J}^s \neq 0$ ) into

$$V Q^s V^{-1} = \left( \frac{A\bar{A}J^s + B\bar{B}\bar{J}^s}{-\bar{A}B(J^s - \bar{J}^s)}, \frac{-AB(J^s - \bar{J}^s)}{A\bar{A}J^s + B\bar{B}\bar{J}^s} \right). \quad (14)$$

This transformed substitution belongs to the cyclic group  $G_{\frac{s+1}{2}, 1}^* \{Q^s\}$  if in  $V$   $AB \neq 0$  and  $\bar{A}\bar{B} = 0$ . We have also  $A\bar{A} + B\bar{B} = 1$ .

Two cases arise.

First case:  $B = \bar{B} = 0$ .  $V = (A, 0; 0, \bar{A} = A^{-1})$  itself belongs to the cyclic group  $G_{\frac{s+1}{2}, 1}^*$ , and, of course, transforms every  $Q^s$  of the group into itself.

Second case:  $A = \bar{A} = 0$ .  $V = (0, B; -\bar{B} = -B^{-1}, 0)$  transforms  $Q^s$  into  $Q^{-s}$ , which is distinct from  $Q^s$  unless  $Q^s$  is of period two.

Within the  $G_{M(n)}$  the cyclic  $G_{\frac{s+1}{2}, 1}^*$  is self-conjugate in exactly the dihedral group  $G_{\frac{s+1}{2}, 1}^*$  composed of the totality of substitutions having the forms  $(A, 0; 0, A^{-1})$ ,  $(0, B; -B^{-1}, 0)$ , where  $A$  and  $B$  are any powers of  $J$ .

Within the  $G_{M(n)}$  a substitution  $Q^s$  is self-conjugate in exactly the  $G_{\frac{s+1}{2}, 1}^*$ , except in the case  $s = q^n = 4k - 1$ , in which the  $Q^{\frac{s+1}{2}}$  is self-conjugate in exactly the dihedral  $G_{\frac{s+1}{2}, 1}^*$ ; this dihedral group is commutative only if it is a four-group ( $s = q^1 = 3$ ).

In the  $G_{M(n)}$  there are  $\left\{ M(s) = \frac{s(s^2 - 1)}{2; 1} \right\} \div 2 \frac{s+1}{2; 1} = \frac{1}{2}s(s-1)$  cyclic groups  $G_{\frac{s+1}{2}, 1}^*$  conjugate to  $G_{\frac{s+1}{2}, 1}^*$ . Each of these is defined by any substitution lying in it (the identity excepted) as the largest cyclic group containing that substitution. The  $\frac{1}{2}s(s-1)$  groups have the identity in common, but otherwise have quite distinct substitutions, of periods divisors of  $\frac{s+1}{2; 1}$ , in all

$$\frac{s \cdot s - 1}{2} \left( \frac{s+1}{2; 1} - 1 \right) = \frac{s(s-1)(s-1; s)}{4; 2}$$

substitutions.

In the  $G_{M(n)}$  there are  $M(s) \div \frac{s+1}{2; 1} = s(s-1)$  substitutions conjugate to  $Q^s$  of which two lie in every  $G_{\frac{s+1}{2}, 1}^*$  (in  $G_{\frac{s+1}{2}, 1}^*$ , for instance, lie  $Q^s$  and  $Q^{-s}$ ). However, for  $s = q^n = 4k - 1$ , there are only  $\frac{1}{2}s(s-1)$  substitutions conjugate to  $Q^{\frac{s+1}{2}}$ , one lying in every  $G_{\frac{s+1}{2}, 1}^*$ .

We have now enumerated all the individual operators of the  $G_{M(\omega)}$ , and likewise all the cyclic and largest commutative<sup>14a</sup> subgroups; for

$$1 + (s^2 - 1) + \frac{s(s+1)(s-3; s-2)}{4; 2} + \frac{s(s-1)(s-1; s)}{4; 2} = \frac{s(s^2-1)}{2; 1} = M(s) .$$

(The identity is not enumerated under I, II, or III, and so is counted separately.)

We find similarly that the  $G_{sM(\omega)}$  ( $q > 2$ ) contains exactly the identity,  $s^2-1$  substitutions of period  $q$ , and substitutions conjugate to the various powers of  $D_p$  (§ 4 (9);  $D_p^2 = P$ ) and of  $D_r$  (§ 5 (16);  $D_r^2 = Q$ ):

$$1 + (s^2 - 1) + \frac{s(s+1)(s-2)}{2} + \frac{s(s-1)(s)}{2} = s(s^2-1) = 2M(s) .$$

It remains to prove the Theorems IV and V.

Obviously, if under a substitution  $V$  of the  $G_{M(\omega)}^{s+1}$  any three marks  $\omega$  are fixed, then  $V$  is the identity  $I$ . The substitution  $V$  ( $V \neq I$ ) may then be separated into three classes according as they leave exactly one mark, two distinct marks, or no mark fixed. These three classes are precisely the three sets I, II, and III.

Under transformation  $(A+D)^2$  is invariant. (The transforming substitution may be any substitution of the  $G_{(s;1)M(\omega)}^{s+1}$ .) To see in how far  $(A+D)^2$  is a characteristic invariant for a system of substitutions conjugate under the  $G_{M(\omega)}^{s+1}$ , we need (by I, II, III) consider only the representative substitutions of such systems

$$(I) \ S_p = (1, \beta; 0, 1) ; \quad (II) \ P^p = (p^p, 0; 0, p^{-p}) ; \quad (III) \ Q^p = (J^p, 0; 0, \bar{J}^p) . \quad (15)$$

These are all of the form

$$V = (A, B; 0, D) \quad (AD=1) . \quad (16)$$

Two particular substitutions  $U_1, U_2$  of the form (16) have the same invariant,

$$(A_1 + D_1)^2 = (A_2 + D_2)^2 , \quad (17)$$

if and only if

$$(A_1, D_1) = (A_2, D_2), (D_2, A_2), (-A_2, -D_2) \text{ or } (-D_2, -A_2) . \quad (18)$$

$U_1$  and  $U_2$  belong then to the same set I, II, III. If they belong to set I, they always have the same invariant 4, but are not always conjugate (Theorem I). If, however, they belong to set II or to set III and have the same invariant, they are always conjugate; this appears from the considerations introduced in the proof of Theorems II and III.

Theorem IV has been proved.

The truth of theorem V appears at once from the consideration of the representative substitutions (15), in the light as to (III) of § 3.

The characterization V(I) holds also in the case  $q = 2$ .

Substitutions of period 2 for  $q$  any are fully characterized by the property  $\alpha + \delta = 0$ ; however for  $q = 2$  the identity also satisfies the condition  $\alpha + \delta = 0$ .

<sup>14a</sup> See the remark in brackets near the beginning of this § 6.

§7. EXTENSION OF THE  $G_s^{(n)}$   $\{S_\beta = (1, \beta; 0, 1)\}$  BY ANY UNIMODULAR SUBSTITUTION  $V = (a, \beta; \gamma, \delta)$  ( $\gamma \neq 0$ ) TO THE  $G_{M(a)}^{s+1}$ , SYSTEMS OF GENERATORS OF THE  $G_{M(a)}^{s+1}$  AND  $G_{(a; 1)M(a)}^{s+1}$

We consider the group  $G\{U\}$  of all binary matrices

$$U = \begin{pmatrix} a, b, \\ c, d, \end{pmatrix} = (a, b, c, d) \quad (ad - bc = 1), \quad (1)$$

whose elements,  $a, b, c, d$  belong to a (finite or infinite) field and are subject only to the condition that  $ad - bc = 1$ .

We introduce

$$T = (0, -1, 1, 0), \quad S_a = (1, a, 0, 1), \quad S_a T = E_a = (a, -1, 1, 0), \quad (2)$$

and have then

$$U_{(d \neq 0)} = E_{(d-1), d^{-1}} E_{-d} E_{-(a+1)d^{-1}}, \quad (3)$$

$$U_{(d=0, \text{ and so } bc=-1)} = E_{t+rb} E_{-a} E_b E_{r+aw}, \quad (4)$$

where in (4)  $r$  is arbitrary and  $t$  and  $w$  are connected only by the equation

$$tc - wb + 1 = a. \quad (5)$$

The group  $G\{S_a\}$  is then extended by  $T$  to the group  $G\{U\}$ .

More generally, the group  $G\{S_a\}$  is extended by any particular matrix  $U = V$  with  $c \neq 0$  to the group  $G\{U\}$ , for

$$T = S_{1-aa^{-1}} V S_{a^{-1}-(a+d)a^{-1}} V S_{1-da^{-1}}. \quad (6)$$

The theorem implied by the caption<sup>15</sup> is an obvious modification of a particular case—for the  $GF[s]$ <sup>16</sup>—of the theorem just proved.

The group  $G\{U\}$  of unitary matrices  $U$  is of course extended by the matrices

$$D_a = (d, 0, 0, 1) \quad (d \neq 0) \quad (7)$$

to the group  $\bar{G}\{U\}$  of all matrices  $U$  of non-vanishing determinant. In view of the relation

$$D_a S_1 D_a^{-1} = S_a \quad (8)$$

the group  $\bar{G}\{U\}$  is generated by the matrices  $D_a$ ,  $S_1$ , and  $T$ .

When the fundamental field is the  $GF[s]$  with the primitive root  $\rho$ , so that  $d = \rho^e$ , the matrix  $D_\rho$  generates the matrices  $D_a$ ,

$$D_a = D_\rho^e, \quad (9)$$

and so the  $\bar{G}\{U\}$  is generated by  $D_\rho$ ,  $S_1$ , and  $T$ . Similarly<sup>17</sup> the group  $\bar{G}\{U\}$  of all matrices  $U$  of determinant a square in the  $GF[s]$  is generated by  $D_\rho$ , and  $S_1$ , and  $T$ .

Hence, for our groups  $G_{M(a)}^{s+1}$ ,  $G_{(a; 1)M(a)}^{s+1}$  we have the respective systems<sup>17</sup> of three generators:

<sup>15</sup> This development of the theorem for the  $G_{M(a)}^{s+1}$  ( $s=q^n$ ) for the general  $n$  involves of necessity essential modifications of the development for  $n=1$  (KLEIN-FRIEDER, *loc. cit.*, pp. 452, 459).

<sup>16</sup> We need also §2, 51'.

<sup>17</sup> BURNSIDE (*loc. cit.*, pp. 117, 118) is the first to exhibit these systems of generators. He works directly with the fractional substitutions. His proof as to the  $G_{M(a)}^{s+1}$  rests upon the lemma of §2, 51' of this paper, of which his proof was inadequate.

$$G_{M(u)}^{s+1}; \quad D_{\rho^2} = P = (\rho, 0; 0, \rho^{-1}), \quad S_1 = (1, 1; 0, 1), \quad T = (0, -1; 1, 0); \quad (10)$$

$$G_{(2;1)M(u)}^{s+1}; \quad D_{\rho} = (\rho, 0; 0, 1), \quad S_1 = (1, 1; 0, 1), \quad T = (0, -1; 1, 0), \quad (11)$$

for the second group is obtained by extending the first by  $D_{\rho}$ , and  $D_{\rho^2} = D_{\rho}^2$ .

§ 8. EVERY  $G_{M(u)}$  ( $s = q^n \neq 2^1, 3^1$ ) IS A SIMPLE GROUP

A self-conjugate subgroup  $G_d$  ( $d > 1$ ) of the  $G_{M(u)}^{s+1}$  must contain all the operators of at least one system of operators conjugate under the  $G_{M(u)}^{s+1}$ .

If the  $G_d$  contain one operator of set I, then it contains all operators conjugate with it and so (§ 6, I) all operators  $S_{\beta}$  where  $\beta$  is either any square or any not-square. Then the  $G_d$  contains all the operators  $S_{\beta}$ ,  $\beta = \text{any}$  (§ 2, 51°), and indeed the  $s + 1$  conjugate groups  $G_s$  (§ 6, I). Then the  $G_d$  is the  $G_{M(u)}$  itself (§ 7), since it contains the  $G_s^*$   $\{S_{\beta}\}$  and an operator  $V = (a, \beta; \gamma, \delta)$  with  $\gamma \neq 0$ .

If the  $G_d$  contain one operator

$$P^{\sigma} = (\rho^{\sigma}, 0; 0, \rho^{-\sigma})$$

of set II, then it contains the conjugate (§ 7, IV) operator

$$V = (\rho^{-\sigma}, \beta; 0, \rho^{\sigma}) \quad (\beta \neq 0),$$

and then the product  $S_{\rho^{\sigma}} = \rho^{\sigma} V$ , which is of set I. Thus the  $G_d$  is the  $G_{M(u)}$  itself.

We now consider a self-conjugate  $G_d$  containing no operators of sets I or II. Except in the cases  $s = 2^1, 3^1$  such a  $G_d$  turns out to be precisely  $G_{d=1}$  = identity itself. Thus the theorem of the caption of § 8 is proved.

We treat this case<sup>18</sup> by the diophantine considerations employed in the Congress paper (pp. 238-42) for the general case. Since

$$M(s) = (2; 1) \cdot s \cdot \frac{s-1}{2; 1} \cdot \frac{s+1}{2; 1}, \quad (1)$$

$d$  a divisor of  $M(s)$  and prime to  $s$  and  $\frac{s-1}{2; 1}$  (by the Cauchy-Sylow theorem) must be a divisor of  $(2; 1) \left( \frac{s+1}{2; 1} \right)$ , so that

$$d = 2^e c_+, \quad c_+ = \left[ \frac{s+1}{2; 1}; d \right], \quad 2^e = 1 \text{ or } 2; 1, \quad (2)$$

where as usual  $[u, v]$  denotes the greatest common divisor of the two positive integers  $u$  and  $v$ . Further (by § 6, III) the  $G_d$  contains and is made up of  $\frac{1}{2}s(s-1)$  cyclic  $G_{d+}$ , one from each of the conjugate cyclic  $G_{\frac{s+1}{2; 1}}$ . Hence

$$d = 1 + \frac{1}{2}s(s-1)(d_+ - 1) = -\frac{s+1}{2; 1} \cdot \frac{s-2}{1; 2} + \frac{1}{2}s(s-1)d_+. \quad (3)$$

<sup>18</sup> BURNAIDE's treatment (*loc. cit.*, pp. 126, 132) of this third case fails either if his  $a + \delta = x^e = 0$  or if  $q = 2$ . As a result of these errors he overlooks the exceptional cases  $s = q^n = 3^1, 2^1$ .



Hence  $d_+$  is a divisor of  $\frac{s+1}{2;1}$  and of  $d$ , and so of  $c_+$ , and further  $c_+$  is a divisor of  $\frac{s+1}{2;1}$  and of  $d$ , and so not of  $\frac{1}{2}s(s-1)$ , and so of  $d_+$ . Hence  $c_+ = d_+$ . Writing  $u$  for  $\frac{1}{2}s(s-1)$ , we have

$$d = 2'd_+ = 1 + u(d_+ - 1) \quad (2' = 1 \text{ or } 2; 1) . \quad (4)$$

Whence, for  $2' = 1$ , we have

$$d = d_+ , (u - 1)(d_+ - 1) = 0 , \quad (5)$$

that is, either the  $G^d = G_{d=d_+=1}$  or  $u = 1$ ,  $s = 2$ ; and, for  $2' = 2$ , we have

$$d = 2d_+ , (u - 2)(d_+ - 1) = 1 , \quad (6)$$

that is,  $d_+ = 2$ ,  $u = 3$ ,  $s = 3$ . Hence, as was stated,  $d = 1$  or  $s = 2$  or  $3$ .

The two remaining groups  $G_{M(u)}^s$ ,  $s = 2, 3$  are in fact composite, being (§ 1) the  $G_s^2$ -symmetric and the  $G_{12}^3$ -alternating with the respective self-conjugate subgroups  $G_2^2$ -cyclic and  $G_4^3$ -four-group.

### § 2. SUBGROUPS OF $G_{M(u)}$ :

THE CYCLIC, THE DIHEDRON, AND THE FOUR-GROUP (WITH CYCLIC BASES NOT OF SET 1)

I recall the formulas for the abstract  $k$ -dihedron group  $G_{2k}$  based on the cyclic  $G^k$ :

$$A^k = I, B^2 = I, AB = BA^{-1} , \quad (1)$$

whence

$$BAB^{-1} = A^{-1}, ABA^{-1} = BA^{-1} . \quad (2)$$

Hence under the  $G_{2k}$  the operators  $A^g$  and  $A^{-g}$  are conjugate ( $g=0, 1, 2, \dots$ ), and the  $k$  operators  $BA^i$  ( $i=0, 1, \dots, k-1$ ) of period two form one set or two sets of conjugate operators according as  $k$  is odd or even, and then the cyclic base  $G_k$  is self-conjugate and the  $k$  cyclic groups  $G_i$  form one set or two sets of conjugate subgroups.

A cyclic  $G_k$  contains one subgroup  $G_d$  for every divisor  $d$  of  $k$ .

A dihedron  $G_{2k}$  contains based on such a cyclic  $G_d$  of its cyclic base  $G_k$  in all  $k/d$  dihedron  $G_{2d}$ . These  $k/d$  groups  $G_{2d}$  form under the  $G_{2k}$  one system or two systems of conjugate subgroups according as  $k/d$  is odd or even.<sup>19</sup> In the former case the  $G_{2d}$  is self-conjugate only under itself. In the latter case the  $G_{2d}$  is self-conjugate under a dihedron  $G_{2,2d}$ .

Cyclic  $G_k$  and dihedron  $G_{2k}$  contain no subgroups other than those just mentioned.

Now (§ 6 II, III) the  $G_{M(u)}$  contains  $\frac{1}{2}s(s \pm 1)$  conjugate largest cyclic subgroups  $G_{\frac{s \mp 1}{2;1}}$  of sets  $\frac{\text{II}}{\text{III}}$ , and hence precisely as many conjugate cyclic  $G_{d_{\mp}}$  where  $d_{\mp}$  is any divisor of  $\frac{s \mp 1}{2;1}$ .

<sup>19</sup> Cf. KLEIN-FRICKER, loc. cit., p. 468.

Every cyclic  $G_{\frac{s+1}{2}, 1}$  lies in one dihedral  $G_{\frac{s+1}{2}, \frac{s+1}{2}}$ ; within the  $G_{(s, 1)M(s)}$  these extend to a cyclic  $G_{s+1}$  and a dihedral  $G_{2(s+1)}$ .

Every cyclic or dihedral subgroup of the  $G_{M(s)}$  with cyclic base not of set I lies in one of these dihedral  $G_{\frac{s+1}{2}, 1}$ .

Every cyclic  $G_{d_{\mp}}$  is the cyclic base of  $\frac{s+1}{2; 1}/d_{\mp}$  dihedral  $G_{2d_{\mp}}$ .

Two dihedral  $G_{2d_{\mp}}$  belonging to the same dihedral  $G_{\frac{s+1}{2}, 1}$  if conjugate under the  $G_{M(s)}$  are conjugate<sup>20</sup> under the  $G_{\frac{s+1}{2}, \frac{s+1}{2}}$ . Under the  $G_{\frac{s+1}{2}, \frac{s+1}{2}}$  the  $\frac{s+1}{2; 1}/d_{\mp}$   $G_{2d_{\mp}}$  form one system or two systems of conjugate groups according as  $\frac{s+1}{2; 1}/d_{\mp}$  is odd or even; within the dihedral  $G_{2(s+1)}$  of the  $G_{(s, 1)M(s)}$  they form one system of conjugate groups.

Every dihedral  $G_{2d_{\mp}}$  ( $d_{\mp} > 2$ ) contains only one basal cyclic  $G_{d_{\mp}}$ .

There are in the  $G_{M(s)}$  in all  $M(s)/2d_{\mp}$  dihedral  $G_{2d_{\mp}}$  ( $d_{\mp} > 2$ ) which form one system or two systems of conjugate groups according as  $\frac{s+1}{2; 1}/d_{\mp}$  is odd or even. In the former case a  $G_{2d_{\mp}}$  is self-conjugate only under itself; in the latter case a  $G_{2d_{\mp}}$  is self-conjugate under a dihedral  $G_{2d_{\mp}}$ . These  $G_{2d_{\mp}}$  within the  $G_{(s, 1)M(s)}$  form only one system.

The  $G_{M(s)}$  always contains operators  $V_s$  of period 2. If  $q = 2$ , then these  $V_s$  all are of set I. If  $q > 2$ , then these  $V_s$  all belong to set II or to set III according as  $\frac{1}{2}(s-1)$  or  $\frac{1}{2}(s+1)$  is even. Now for  $q > 2$ ,  $s = q^n$  is of the form  $4h+1$  or  $4h-1$ , according as the Jacobi-Legendre symbol  $\left(\frac{-1}{s}\right)$  is  $+1$  or  $-1$ ; hence  $\frac{1}{2}\left[s - \left(\frac{-1}{s}\right)\right]$  is even, and all the  $V_s$  belong to the  $\frac{1}{2}\left[s + \left(\frac{-1}{s}\right)\right]$  conjugate cyclic  $G_{\frac{1}{2}\left[s - \left(\frac{-1}{s}\right)\right]}$  (§ 6 II, III).

The four-groups  $G_4 = G_{2d}$  ( $d = 2$ ) with cyclic bases not of set I remain for consideration. The  $G_{M(s)}$  ( $s = q^n$ ,  $q > 2$ ) contains (§ 6 II, III)  $\frac{1}{2}s\left[s + \left(\frac{-1}{s}\right)\right]$  conjugate cyclic  $G_4$ . Every  $G_4$  lies in  $\frac{1}{4}\left[s - \left(\frac{-1}{s}\right)\right]$  four-groups  $G_4$ . Every  $G_4$  contains 3  $G_2$ . There are in the  $G_{M(s)}$  ( $s = q^n$ ,  $q > 2$ ) in all  $M(s)/12$  four-groups  $G_4$ ; these form one system or two systems of conjugate groups according as  $\frac{1}{4}\left[s - \left(\frac{-1}{s}\right)\right]$  is odd or even,

<sup>20</sup>[Proof for  $d_{\mp}=2$  (and so  $q>2$ ). A four-group  $G_4$  ( $d_{\mp}=2$ ;  $q>2$ ) has three substitutions  $T_1, T_2, T_3$  of period 2. Within the  $G_{M(s)}$  the  $G_4$  is self-conjugate under a group  $G_{4u} \{U\}$ . The substitutions  $U$  as transformers, I affirm, permute the  $T$ 's transitively; from this fact the statement of the text follows at once. For otherwise one  $T$ , say  $T_1$ , would be self-conjugate under the  $G_{4u}$ , and so the  $G_{4u}$  would be a subgroup (in fact, dihedral) of the dihedral

$G_{\frac{s+1}{2}, \frac{s+1}{2}}$  corresponding to  $T_1$  (§ 6, II, III); hence the  $G_4$  would be self-conjugate under exactly itself ( $u=1$ ) or a dihedral  $G_4$  ( $u=2$ ), and so would be within the  $G_{M(s)}$  one of a system of  $M(s)/4$  or  $M(s)/8$  conjugate four-groups; but, as is immediately proved in the text, the  $G_{M(s)}$  ( $q>2$ ) contains in all  $M(s)/12$  four-groups.

that is, according as  $s = q^n$  has the form  $8h \pm 3$  or  $8h \pm 1$ . In the former case a  $G_i$  is self-conjugate under a  $G_{12}$ , in the latter case under a  $G_{24}$ . These  $G_i$  form within the  $G_{2M(u)}$  only one system; every  $G_i$  is self-conjugate under a  $G_{24}$ . These groups  $G_{12}$ ,  $G_{24}$  have the tetrahedron, octahedron type (§§ 13 ( $4^\circ$ ), 15 (IV, V), 18).

We have now completely enumerated the cyclic, dihedron, and four-group subgroups of the  $G_{M(u)}$  whose cyclic bases are not of set I.

Within the  $G_{2M(u)}$  ( $q > 2$ ) every cyclic  $G_{d\mp}$  ( $d_{\pm} > 1$ ) of the cyclic  $G_{s\mp}$  of the  $G_{M(u)}$  extends to a definite cyclic  $G_{2d\mp}$  of the cyclic  $G_{s\mp}$  and every dihedron  $G_{2d\mp}$  of the  $G_{M(u)}$  to a definite dihedron  $G_{2,2d\mp}$  (which belong to the  $G_{M(u)}$  only if  $(s \mp 1)/2d_{\mp}$  is even). Conversely, if  $d$  is even and  $> 2$ , every cyclic  $G_d$  and dihedron  $G_{2,d}$  of the  $G_{2M(u)}$  (with cyclic base not of set I) is the extension of a definite similar group of the  $G_{M(u)}$ . We find that:

The  $G_{2M(u)}$  ( $q > 2$ ) contains for every divisor  $d > 2$  of  $s \pm 1$  one system of conjugate cyclic  $G_d$  and one or two systems of conjugate dihedron  $G_{2,d}$  according as  $d$  is even or odd. The  $G_{2M(u)}$  contains a system of  $\frac{1}{2}s \left[ s + \left( \frac{-1}{s} \right) \right]$  conjugate  $G'_i$  (of the  $G_{M(u)}$ ) and another system of  $\frac{1}{2}s \left[ s - \left( \frac{-1}{s} \right) \right]$  conjugate  $G'_i$  (not of the  $G_{M(u)}$ ). The  $G_{2M(u)}$  contains a system of  $M(s)/12$  conjugate four-groups  $G'_{2,2}$  and another system of  $M(s)/4$  four-groups  $G'_{2,2}$ ; a  $G'_{2,2}$  contains three  $G'_i$ , while a  $G'_{2,2}$  contains one  $G'_i$  and two  $G'_i$ .

#### §10. SUBGROUPS OF $G_{M(u)}$ :

##### COMMUTATIVE SUBGROUPS OF THE $s+1$ $G'_i$ OF SET I (§6, I)

The  $s+1$   $G'_i$  of set I are conjugate under the  $G_{M(u)}$  (§6, I). We study the subgroups of  $G'_i$   $\{S_\beta\}$ . This is a commutative group of order  $s = q^n$ ; every operator is of period  $q$ . To every subgroup  $G_{q^m}$  of order  $q^m$  ( $m \leq n$ ) corresponds an additive-group of rank  $m$  in the Galois-field  $GF[s = q^n]$ , and conversely.

Writing  $B(n, m, x)$  for Gauss's cyclotomic function<sup>21</sup> ( $n, m$ ),

$$B(n, m, x) = \frac{(x^n - 1)(x^{n-1} - 1) \dots (x^{n-m+1} - 1)}{(x^m - 1)(x^{m-1} - 1) \dots (x - 1)}, \quad (1)$$

we readily see that there are in the  $G_{s=q^n}$  exactly  $B(n, m, q)$  subgroups  $G_{q^m}$ .

We consider one such subgroup  $G = G_{q^m} \{S_\lambda\}$  where  $\lambda$  ranges over an additive-group  $[\lambda_1, \dots, \lambda_m]$  of rank  $m$  (§2, 19°) and inquire as to the largest group  $H$  within the  $G_{M(u)}$  under which  $G$  is self-conjugate.  $H$  must be a subgroup of the  $G'_{s=(s-1)} \{V = (a, 0; 0, a^{-1})\}$  (§6, I).  $V$  transforms  $S_\lambda$  into  $S_{a^{-1}\lambda}$ , and the group  $G_{q^m} \{S_\lambda\}$ ,  $\lambda = [\lambda_1, \dots, \lambda_m]$ , into the group  $G' = G_{q^m} \{S_{a^{-1}\lambda}\} = G_{q^m} \{S_\lambda\}$ ,  $\lambda' = [a^2\lambda_1, \dots, a^2\lambda_m]$ . Suppose that there are in all  $e$  marks  $a = \epsilon$  (say  $\epsilon_1, \dots, \epsilon_e$ ), such that the two

<sup>21</sup> GAUSS, *Werke*, Vol. II, p. 16; GAUSS-MASER, p. 467.

additive-groups  $[\lambda_1, \dots, \lambda_m]$  and  $[\epsilon^2 \lambda_1, \dots, \epsilon^2 \lambda_m]$  are the same. (The marks  $\epsilon = \pm 1$ ; 1 are such marks  $\epsilon$ . Hence  $e \geq 2; 1$ ). Then the  $\frac{se}{2; 1}$  operators  $V$ :

$$V = (\epsilon, \epsilon\beta; 0, \epsilon^{-1}) = (\epsilon^2, 0; 0, 1) (1, \beta; 0, 1), \quad (2)$$

where independently  $\beta$  ranges over the  $GF[s]$  and  $\epsilon$  ranges over the  $e$  marks  $\epsilon_1, \dots, \epsilon_e$ , constitute the largest group  $H = G_{\frac{se}{2; 1}}$  under which  $G = G_{q^m} \{S_\lambda\}$  is self-conjugate.

The determination for any given group  $G = G_{q^m} \{S_\lambda\}$  of the group  $H$  has then been made to depend upon the determinations for the additive-group  $[\lambda_1, \dots, \lambda_m]$  of rank  $m$ , *first*, say of its *multipliers*  $\kappa$ , that is, of all the marks  $\kappa$  such that

$$[\lambda_1, \dots, \lambda_m] = [\kappa \lambda_1, \dots, \kappa \lambda_m], \quad (3)$$

and, *secondly*, of all the marks  $\epsilon$  whose squares are multipliers.

It turns out (§11) that the multipliers  $\kappa$  are the marks  $\kappa$  ( $\kappa \neq 0$ ) of (a definite so-called) the multiplier Galois-field  $GF[q^k]$  of the additive-group  $[\lambda_1, \dots, \lambda_m]$ . We discriminate the cases

$$(q > 2) \ n/k \text{ even, } n/k \text{ odd; } (q = 2). \quad (4)$$

Then (§ 2, 50°, 39°) there are

$$e = (2, 1; 1) (q^k - 1) \quad (5)$$

marks  $\epsilon$ , and the group  $G = G_{q^m} \{S_\lambda\}$  is within the  $G_{M^{(n)}}$  or the  $G_{\frac{s(q^k-1)}{2; 1}}^{(n)}$  self-conjugate under the group  $H$  (2) of order

$$\frac{se}{2; 1} = \frac{s(q^k - 1)}{1, 2; 1}, \quad (6)$$

and so is one of a system of

$$\frac{(s+1)(q^n - 1)}{(2, 1; 1)(q^k - 1)} \text{ or } \frac{q^n - 1}{(2, 1; 1)(q^k - 1)} \quad (7)$$

conjugate groups.

Within the  $G_{M^{(n)}} (q > 2)$  the subgroups  $G_{q^m} \{S_\lambda\}$  of the type of § 10 are precisely those of the  $G_{M^{(n)}}$ . However, the systems of conjugate groups of the  $G_{M^{(n)}}$  with odd  $n/k$  (4) unite by pairs to form the systems of conjugate groups of the  $G_{M^{(n)}}$ , for the two  $G_{q^m} \{S_\lambda\}$  with  $\lambda = [\lambda'_1, \dots, \lambda'_m]$  and  $\lambda = [\rho' \lambda'_1, \dots, \rho' \lambda'_m]$ , where  $\rho'$  is a primitive root of the multiplier  $GF[q^k]$  of the additive-group  $[\lambda'_1, \dots, \lambda'_m]$ , are conjugate in the  $G_{M^{(n)}}$  but not in the  $G_{M^{(n)}}$ .—The group  $G_{q^m} \{S_\lambda\}$  is within the  $G_{(2; 1)M^{(n)}}$  self-conjugate under the group  $H'_{q^n (q^k - 1)}$  of substitutions  $D_\kappa S_\beta$  ( $\kappa \neq 0$ )

$$(\kappa, \kappa\beta; 0, 1) = (\kappa, 0; 0, 1) (1, \beta; 0, 1). \quad (8)$$

#### § 11. ADDITIVE-GROUPS IN THE GALOIS-FIELD $GF[q^n]$ AND THEIR CORRESPONDING MULTIPLIER GALOIS-FIELDS

In the  $GF[s = q^n]$  the additive-group  $[\lambda_1, \dots, \lambda_m]$  of rank  $m$  ( $m \leq n$ ) based upon the  $m$  marks  $\lambda_i$  linearly independent with respect to the  $GF[q]$  contains (§ 2, 19°) the  $q^m$  distinct marks  $\lambda$  of the form

$$\lambda = c_1 \lambda_1 + \dots + c_m \lambda_m \quad (1)$$

where the  $c_i$ 's are integral marks.

If the  $q^m$  marks  $\lambda$  are multiplied by any particular mark  $\mu$  ( $\mu \neq 0$ ) then the resulting  $q^m$  marks  $\mu\lambda$  constitute the additive-group  $[\mu\lambda_1, \dots, \mu\lambda_m]$  or  $\mu[\lambda_1, \dots, \lambda_m]$  likewise of rank  $m$ . We may say that  $[\mu\lambda_1, \dots, \mu\lambda_m]$  is derived from  $[\lambda_1, \dots, \lambda_m]$  by multiplication by  $\mu$ .

We seek (§ 10, (3)) the multipliers  $\mu = \kappa$  of the  $[\lambda_1, \dots, \lambda_m]$ , that is, those multipliers  $\mu = \kappa$  under multiplication by which  $[\lambda_1, \dots, \lambda_m]$  is invariant. These multipliers  $\kappa$  constitute obviously a multiplicative-group. Further, together with the mark 0, they constitute an additive-group with respect to the field  $GF[q]$  of integral marks. For, if  $\kappa_1$  and  $\kappa_2$  are any two multipliers  $\kappa$  then under multiplication by  $\mu = \kappa_1 + \kappa_2$  the  $[\lambda_1, \dots, \lambda_m]$  becomes  $[\mu\lambda_1, \dots, \mu\lambda_m]$  which is an additive-group, on the one hand, included within the  $[\lambda_1, \dots, \lambda_m]$  and, on the other hand, unless  $\mu = 0$ , of rank  $m$ . Hence, the mark  $\mu$  ( $\mu = \kappa_1 + \kappa_2$ ) is either 0 or a multiplier  $\kappa$ .

The marks 0 and the multipliers  $\kappa$  constitute then a Galois-field  $GF[q^k]$  of rank  $k$  included within the fundamental  $GF[q^n]$ . We call this  $GF[q^k]$  *the multiplier Galois-field*, and every  $GF[q^{k'}]$  included within it a *multiplier Galois-field* of the additive-group  $[\lambda_1, \dots, \lambda_m]$  of rank  $m$ , and the integer  $k$  its *multiplicative-index*.  $k'$  is a divisor of  $k$ , and  $k$  is a divisor of  $n$  (§ 2, 39°) and indeed (as we shall prove in a moment) also of  $m$ . There are  $q^k - 1$  multipliers  $\kappa$ .

The multiplier  $GF[q^k]$  of the additive-group  $[\lambda_1, \dots, \lambda_m]$  is the (largest) additive-group common to the  $m$  additive-groups

$$[\lambda_i^{-1}\lambda_1, \dots, \lambda_i^{-1}\lambda_m] \quad (i=1, 2, \dots, m), \quad (2)$$

and it is, moreover, contained in the  $q^m - 1$  additive-groups

$$[\lambda^{-1}\lambda_1, \dots, \lambda^{-1}\lambda_m] \quad (\lambda \text{ of } [\lambda_1, \dots, \lambda_m], \lambda \neq 0). \quad (3)$$

The additive-group  $[\lambda_1, \dots, \lambda_m]$  of rank  $m$  with respect to the  $GF[q]$  may with respect to any such multiplier  $GF[q^{k'}]$  be exhibited as an additive-group  $[\lambda'_1, \dots, \lambda'_{m/k'} \mid GF[q^{k'}]]$  of rank  $m/k'$  (§ 2, 24°). Hence, every such  $k'$  is a divisor of  $m$ , and the multiplicative-index  $k$  is itself such a  $k'$ .

Conversely, every additive-group with respect to any  $GF[q^k]$  has that  $GF[q^k]$  as a multiplier  $GF[q^k]$ .

With respect to any four positive integers  $q$  (a prime),  $n, m, k$ , we introduce the arithmetic-functional notations

$$A(q, n, m, k), \quad A[q, n, m, k] \quad (3)$$

to designate the number within a  $GF[q^n]$  of additive-groups of rank  $m$  having the included  $GF[q^k]$  as respectively  $a$ , the multiplier  $GF[q^k]$ . We set at once

$$A(q, n, m, k) = A[q, n, m, k] = 0, \quad (4)$$

unless the  $q, n, m, k$  satisfy the conditions:

$$q \text{ a prime, } n \geq m, k \text{ a common divisor of } n \text{ and } m. \quad (5)$$

Attending then only to cases (5) we introduce the integers  $h, l, u$ , and  $v$ , so that

$$[m, n] = l = hk, m = lu = hku, n = lv = hkv, [u, v] = 1, u \leq v. \quad (6)$$

Then we find easily

$$A(q, n, m, k) = \frac{(q^n - 1)(q^n - q^k)(q^n - q^{2k}) \cdots (q^n - q^{(hu-1)k})}{(q^m - 1)(q^m - q^k)(q^m - q^{2k}) \cdots (q^m - q^{(lu-1)k})}. \quad (7)$$

We recall the cyclotomic function  $(\lambda, \mu)$  of Gauss with its more explicit notation  $B(\lambda, \mu, x)$  of §10 (1). In this notation we have

$$A(q, n, m, k) = B(hv, hu, q^k) = B(n/k, m/k, q^k). \quad (8)$$

For fixed  $q, n$ , and  $m$ , and so for fixed  $l, u$ , and  $v$  we write, for variable  $k$  and  $h$ , complementary divisors of  $l$ , ( $l = hk$ ),

$$A(q, n, m, k) = B(h), \quad A[q, n, m, k] = B[h], \quad [m, n] = l = hk. \quad (9)$$

The functions  $B(h)$ ,  $B[h]$  are well-defined, single-valued arithmetic functions of the argument  $h$  for all positive integers  $h$  which are divisors of  $l = [m, n]$ .

We have obviously

$$B(h) = \sum_{d|h} B[d], \quad (10)$$

where the summation-remark  $d|h$  denotes that  $d$  runs over the divisors of  $h$ . Hence, introducing Kronecker's<sup>22</sup> symbols  $\delta_{ij}$  and  $\epsilon_t$  for all numbers  $i, j$  and for all positive integers  $t$ ,

$$\delta_{ij} = \begin{cases} 1, & \text{if } i=j, \\ 0, & \text{if } i \neq j; \end{cases} \quad (11)$$

$$\epsilon_t = \begin{cases} 1, & \text{if } t \text{ is } 1, \\ 0, & \text{if } t \text{ is divisible by the square of any prime,} \\ (-1)^r, & \text{if } t \text{ is the product of } r \text{ distinct primes,} \end{cases} \quad (12)$$

and noticing that

$$\sum_{d|t} \epsilon_d = \delta_{1t}, \quad (13)$$

we find<sup>22, 23</sup> that

$$B[h] = \sum_{\substack{d|h \\ dd'=h}} \epsilon_d B[d']. \quad (14)$$

From (9, 6, 14, 8) we have the determination

$$A[q, n, m, k] = \sum_{d|h} \epsilon_d A(q, n, m, kd) = \sum_{d|h} \epsilon_d B(n/kd, m/kd, q^{kd}) \quad ([m, n] = l = hk). \quad (15)$$

We are thus led to associate with Gauss's function  $B(hv, hu, x)$ , where  $h, u, v$  are positive integers,  $v \geq u$ , and  $[u, v] = 1$ , the (so far as I know) new cyclotomic function  $B[hv, hu, x]$ , so that

$$B[hv, hu, x] = \sum_{d|h} \epsilon_d B(hv/d, hu/d, x^d), \quad (16)$$

$$B(hv, hu, x) = \sum_{d|h} B[hv/d, hu/d, x^d]. \quad (17)$$

The functions  $B(hv, hu, x)$ ,  $B[hv, hu, x]$  are rational integral functions of  $x$  with rational integral coefficients. Further  $B[hv, hu, 0] = \delta_{1h}$ , and it may be shown that  $B[hv, hu, x]$  has the factor  $x^{1-\delta_{1h}}(x^{hv}-1)/(x-1)$ .

<sup>22</sup> Cf. KRONECKER, *Berliner Sitzungsberichte*, July 29, 1886, p. 708, and April 5, 1888, p. 418.

VII, *Über einige Sätze aus der Theorie der Kreistheilung*, pp. 360-70.

<sup>23</sup> DIRICHLET-DEDEKIND, *Zahlentheorie*, 4th ed., supp.

We have, then, in the notations of § 10 (1), § 11 (3, 5, 6), from (8, 15),

$$A(q, n, m, k) = B(hv, hu, q^k), \quad (18)$$

$$A[q, n, m, k] = B[hv, hu, q^k]. \quad (19)$$

#### § 12. SUBGROUPS OF $G_{M(s)}$ :

##### NON-COMMUTATIVE SUBGROUPS OF THE $s+1$ $G_{\frac{s(s-1)}{2;1}}^{(\infty)}$ (§ 6, I)

(1°) The  $s+1$   $G_{\frac{s(s-1)}{2;1}}^{(\infty)}$  are conjugate under the  $G_{M(s)}$  (§ 6, I). We study the  $G_{\frac{s(s-1)}{2;1}}^{(\infty)} \{(\alpha, \beta; 0, \alpha^{-1})\}$ . This is made up (§ 6 I, II) of the  $G_s^{(\infty)} \{S_\beta\}$  as self-conjugate subgroup and the  $s$  conjugate groups  $G_{\frac{s-1}{2;1}}^{(\infty \mu)}$  ( $\mu \neq \infty$ ). We make the ordinary matricular arrangement of the substitutions of  $G_{\frac{s(s-1)}{2;1}}^{(\infty)}$  with respect to the (self-conjugate) subgroup  $G_s^{(\infty)}$ ; we may take the substitutions of the cyclic  $G_{\frac{s-1}{2;1}}^{(\infty)}$  as the extending substitutions; hence the (abstract) quotient-group  $G_{\frac{s(s-1)}{2;1}}^{(\infty)} / G_s^{(\infty)}$  is the (abstract) cyclic  $G_{\frac{s-1}{2;1}}$ ; its only subgroups are the cyclic  $G_{d_-}$  for every  $d_-$  divisor of  $\frac{s-1}{2;1}$ .

Every subgroup  $G^*$  of the  $G_{\frac{s(s-1)}{2;1}}^{(\infty)}$  has its substitutions lying on the face of that matricular arrangement in such a way as to exhibit their matricular arrangement with respect to the subgroup  $G = G_{q^m} \{S_\lambda\}$ ,  $\lambda = [\lambda_1, \dots, \lambda_m]$ , of its substitutions of period  $q$  (§ 10). This is a self-conjugate subgroup. The quotient-group  $G^* / G_{q^m}$  must be such a cyclic  $G_{d_-}$ . A substitution  $V$  in that line of the  $G^*$ -arrangement which corresponds to the generator of the quotient cyclic  $G_{d_-}$  has itself the period  $d_-$ , for  $V^{d_-}$  lies in the group  $G_{q^m}$ .

The  $s$  groups  $G_{\frac{s-1}{2;1}}^{(\infty \mu)}$  are conjugate (§ 6, II) under the transformers  $S_\beta$ , which transform  $G_{q^m} \{S_\lambda\}$  into itself.

We obtain a representative under  $G_{\frac{s(s-1)}{2;1}}^{(\infty)}$  of every  $G^*$  based on a particular  $G_{q^m} \{S_\lambda\}$  when we extend the  $G_{q^m} \{S_\lambda\}$  by the various cyclic  $G_{d_-}^{(\infty)}$  under which it is self-conjugate.

Now the  $G_{q^m} \{S_\lambda\}$  is self-conjugate under the group  $H_{\frac{s(q^k-1)}{1;2;1}}^{(\infty)}$  (§ 10 (2, 4, 5, 6)) obtained by extending the  $G_s^{(\infty)}$  by the cyclic  $G_{\frac{q^k-1}{1;2;1}}^{(\infty)}$ . Our representative  $G_{q^m d_-}^*$  is obtained by extending the  $G_{q^m} \{S_\lambda\}$  by a cyclic  $G_{d_-}^{(\infty)}$  subgroup of this  $G_{\frac{s(q^k-1)}{1;2;1}}^{(\infty)}$ . The  $G_{q^m d_-}^*$  is made up of the  $G_{q^m} \{S_\lambda\}$  and  $q^m$  cyclic  $G_{d_-}^{(\infty)}$ ; it is self-conjugate under the  $G_{\frac{s(q^k-1)}{1;2;1}}^{(\infty)}$  obtained by extending the  $G_{q^m} \{S_\lambda\}$  by the  $G_{\frac{q^k-1}{1;2;1}}^{(\infty)}$ , and so is one of a system of  $(q^s + 1) q^{s-m} (q^s - 1) / (2, 1; 1) (q^k - 1)$  conjugate groups.

A subgroup  $G^*$  of  $G_{\frac{s(s-1)}{2;1}}^{(\infty)}$  which contains no substitution of period  $q$  is a cyclic subgroup  $G_{d_-}^{(\infty)}$  of one of the cyclic  $G_{\frac{s-1}{2;1}}^{(\infty \mu)}$ . This remark is of importance in the theory of those subgroups of the  $G_{M(s)}$  which contain no operators of period  $q$ .

(2°) The typical non-commutative subgroup of the  $G_{s(s-1)}^{(\infty)}$  of the  $G_{(s-1)M(s)}$  is the  $G_{q^m d_-}^*$  obtained by extending the general  $G_{q^m}\{S_\lambda\}$  of § 10 by a cyclic  $G_{d_-}^{(\infty)}$  subgroup of the  $G_{q^{k-1}}^{(\infty)}\{D_\kappa\}$  ( $\kappa \neq 0$ ) (§ 10 (8)). This  $G_{q^m d_-}^*$  is self-conjugate under the  $G_{q^m(q^k-1)}\{D_\kappa S_\lambda\}$ , and thus determines a system of

$$(q^n + 1)q^{n-m}(q^n - 1)/(q^k - 1)$$

conjugate subgroups.

(3°) The  $G_{s(s-1)}^{(\infty)}$  contains no octahedron or icosahedron subgroups, and tetrahedron subgroups only in case  $q = 2$ ,  $n = \text{even}$ ,  $m = k = 2$ ,  $e = 3$ . In this case the  $G_{(s-1)M(s)} = G_{M(s)}$  ( $q = 2$ ) contains  $M(s)/12$  tetrahedron groups, all conjugate with the group  $G_{s,3=12}^{(\infty)}(a, \beta; 0, 1)$ , where  $a$  ( $a \neq 0$ ) and  $\beta$  are marks of the  $GF[2^3]$ . (Cf. § 15, 2°.)

### § 13. SUBGROUPS OF $G_{M(s)}$ CONTAINING OPERATORS OF PERIOD $q$

(1°) The substitutions of period  $q$ , hence of set I, of the (any) subgroup  $G_\Omega$  of the  $G_{M(s)}^{s+1}$  distribute themselves over certain  $s+1$  subgroups  $G_{q^m \mu}^{(\mu)}$  of the  $s+1$  groups  $G_{q^\mu}^{(\mu)}$  (§ 10). Of the  $s+1$  orders  $q^{m\mu}$  at least one is by hypothesis greater than 1. By suitable transformation within the  $G_{M(s)}$  we arrange it so that  $q^{m\mu} > 1$ ,  $m = m_\mu > 0$ . Under the  $q^m$  transformers  $S_\beta$  of the  $G_{q^\mu}^{(\mu)}$  the remaining  $G_{q^m \mu}^{(\mu)}$  with  $m_\mu > 0$  ( $\mu \neq \infty$ ) (if any) arrange themselves in sets, each consisting of  $q^m$  conjugate groups. Under the  $G_\Omega$  the  $G_{q^m \mu}^{(\mu)}$  is then one of a set of  $1 + f q^m$  conjugate groups, where  $f$  is a positive or zero integer. The  $G_\Omega$  contains no group  $G_{q^m \mu}^{(\mu)}$  ( $m_\mu > 0$ ) other than the  $1 + f q^m$  groups of this set. For any such group  $G_{q^m \mu}^{(\mu)}$  would be one of a set of  $n_\mu$  conjugate groups, where  $n_\mu$  would necessarily have at the same time the forms (by parity of reasoning)  $n_\mu = 1 + f_\mu q^{m_\mu}$  and (by the previous considerations)  $n_\mu = f'_\mu q^m$ .

Every  $G_\Omega$  which contains operators of period  $q$  contains these operators in  $1 + f q^m$  groups  $G_{q^m \mu}^{(\mu)}$  conjugate under the  $G_\Omega$ , where for each group  $G_\Omega$   $f$  and  $m$  are properly determined integers  $f \geq 0$ ,  $m > 0$ .

The groups  $G_\Omega$  with  $f = 0$  we have already enumerated (§§ 10, 12), since every such group lies in a group  $G_{s(s-1)}^{(\mu)}$  (§ 6, I).

(2°) We proceed to the general group  $G_\Omega$  with  $f \geq 1$  and with  $m = m_\mu > 0$ . It contains  $1 + f q^m$  groups  $G_{q^m \mu}^{(\mu)}$  conjugate with a certain group  $G_{q^m}^{(\infty)} = G_{q^m}\{S_\lambda\}$ ,  $\lambda = [\lambda_1, \dots, \lambda_m]$ . This  $G_{q^m}^{(\infty)}$  is within the  $G_\Omega$  self-conjugate under a certain  $G_{q^m d_-}^{(\infty)}$  (§ 12), and hence

$$\Omega = (1 + f q^m) q^m d_- . \quad (1)$$

By transformation<sup>24</sup> by a substitution of the  $G_{s(s-1)}^{(\infty)}$  suitably chosen we arrange it so that this  $G_{q^m d_-}^{(\infty)}$  is

$$G_{q^m d_-}^{(\infty)}\{V_{\eta, \lambda}\} , \quad V_{\eta, \lambda} = \begin{pmatrix} \eta & \eta \lambda \\ 0 & \eta^{-1} \end{pmatrix} = \begin{pmatrix} \eta & 0 \\ 0 & \eta^{-1} \end{pmatrix} \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} = P_\eta S_\lambda , \quad (2)$$

<sup>24</sup> If  $d_- = 1$ , this transformation is not needed here and remains available for other purposes.



obtained by extending the  $G_{q^m}^{(\infty)} \{S_\lambda\}$  by the cyclic  $G_{d_-}^{(\infty)} \{P_\eta\}$  contained within the cyclic  $G_{\frac{q^m}{2}}^{(\infty)} \{P_\epsilon\}$ . There are  $(2; 1)d_-$  marks  $\eta$ , the distinct powers of a mark  $\eta_0$ ; in particular,  $\eta_0^{d_-} = -1$ . Of course,  $P_\eta = P_{-\eta}$ ,  $V_{\eta, \lambda} = V_{-\eta, \lambda}$ .

Now (§§ 10, 11) the squares of the marks  $\eta$  (*qud* marks  $\epsilon$ ) are multipliers  $\kappa$  of the  $[\lambda_1, \dots, \lambda_m]$ . The multiplier  $GF[q^k]$  lies in  $[\lambda_0^{-1} \lambda_1, \dots, \lambda_0^{-1} \lambda_m]$  where  $\lambda_0$  is any mark  $\lambda \neq 0$  (§ 11 (3)). We have

$$P_\sigma(\alpha, \beta; \gamma, \delta) P_\sigma^{-1} = (\alpha, \sigma^2 \beta; \sigma^{-2} \gamma, \delta), \quad P_\sigma V_{\eta, \lambda} P_\sigma^{-1} = V_{\eta, \sigma^2 \lambda}. \quad (3)$$

Hence, by transformation of  $G_{M(\omega)}$  and  $G_\Omega$  by  $P_{\lambda_0^{-1}} = D_{\lambda_0^{-1}}(\omega' = \lambda_0^{-1} \omega)$  of the  $G_{(2; 1) M(\omega)}$  (§ 4) we have  $G_{M(\omega)}$  and within it  $G_\Omega^*$  whose additive-group  $[\lambda_1^*, \dots, \lambda_m^*]$  contains the mark  $\lambda^* = 1$  and so the marks of its multiplier  $GF[q^k]$ . We suppose that this transformation has been made; we drop the  $*$ 's.

(3°) The  $G_\Omega$  of order  $\Omega = (1 + fq^m) q^m d_-$  is obtained by extending the  $G_{q^m d_-}^{(\infty)} \{V_{\eta, \lambda}\}$  by the identity and certain  $f q^m$  extenders  $V_j$  ( $j=1, \dots, f q^m$ ),

$$V_j = \left( \frac{\alpha_j, \beta_j}{\gamma_j, \delta_j} \right), \quad V_j V_{\eta, \lambda} = \left( \frac{\alpha_j \eta, \alpha_j \eta \lambda + \beta_j \eta^{-1}}{\gamma_j \eta, \gamma_j \eta \lambda + \delta_j \eta^{-1}} \right) \quad (\eta_j \neq 0; j=1, \dots, f q^m). \quad (4)$$

The  $G_\Omega$  contains  $(1 + fq^m)(q^m - 1)$  substitutions of set I. Of these  $q^m - 1$  are in the  $G_{q^m}^{(\infty)} \{S_\lambda\}$ . The remaining  $f q^m (q^m - 1)$  are substitutions  $V_j V_{\eta, \lambda}$  satisfying the necessary and sufficient condition (§ 6, V(I)):

$$\alpha_j \eta + \delta_j \eta^{-1} + \gamma_j \eta \lambda = \pm 2. \quad (5)$$

For given  $V_j$  and  $\eta$  ( $\gamma_j \neq 0, \eta \neq 0$ ) there are at most  $2; 1$  values of  $\lambda$  satisfying (5). For given  $V_j$  ( $\gamma_j \neq 0$ ) there are at most  $(2; 1)d_-$  substitutions  $V_{\eta, \lambda}$  ( $V_{\eta, \lambda} = V_{-\eta, \lambda}$ ) such that  $V_j V_{\eta, \lambda}$  is of set I. Hence by §§ 10, 11, 12

$$q^m - 1 \leq (2; 1)d_- \leq e = (2, 1; 1)(q^k - 1) \leq (2, 1; 1)(q^m - 1). \quad (6)$$

There are in all two cases:

$$[A] \quad m = k, \quad q^k - 1 = (2; 1)d_-, \quad \Omega = (1 + fq^k) q^k (q^k - 1) / (2; 1).$$

$$[B] \quad m = k, \quad q > 2, n/k = \text{even}, \quad q^k - 1 = d_-, \quad \Omega = (1 + fq^k) q^k (q^k - 1).$$

In each case  $m = k$  and so (2°) the additive-group  $[\lambda_1, \dots, \lambda_m]$  is its own multiplier  $GF[q^k]$  and every  $\lambda$  is 0 or a multiplier  $\kappa$ .

(4°) [Case A.]  $q^k - 1 = (2; 1)d_-$  and the  $\eta$ 's are the  $\kappa$ 's. The equation (5) with  $V_j$  and  $\eta$  (or  $\kappa$ ) given has exactly  $2; 1$  solutions  $\lambda$ . We may replace  $V_j$  by a  $V_j V_{\kappa, \lambda}$  of set I. We set then  $\delta_j = 2 - \alpha_j$ ,  $\eta = \kappa$ , and determine by the equation

$$\alpha_j (\kappa - \kappa^{-1}) + \gamma_j \lambda_j \kappa \pm = \pm 2 - 2 \kappa^{-1} \quad (7)$$

for every  $j$  and  $\kappa$  and sign  $\pm$  a mark  $\lambda$ , viz.,  $\lambda_j \kappa \pm$  — a mark  $\eta \lambda$  being a mark  $\lambda$ .

*Theorem:* The  $G_\Omega$  is<sup>25</sup> the group  $G_{M(q^k)}$  of  $\Omega = M(q^k) = (1 + q^k) q^k (q^k - 1) / (2; 1)$  substitutions

$$V = (\lambda_1, \lambda_2; \lambda_3, \lambda_4) \quad (\lambda_1 \lambda_4 - \lambda_2 \lambda_3 = 1) \quad (8)$$

where the  $\lambda_1, \lambda_2, \lambda_3, \lambda_4$  belong to the  $GF[q^k]$ .

<sup>25</sup> It appears below, however, that for  $q^k = 2^1$  there are other groups  $G_\Omega$ . For  $q^k = 3, 5, 2^1$  the  $G_{M(q^k)}$  is of type tetrahedron, icosahedron, icosahedron (§ 1).

We shall prove that every  $\alpha_j, \beta_j, \gamma_j, \delta_j$  is a  $\lambda$ . Hence  $G_\Omega$  is a subgroup of the  $G_{M(q^k)}$  (8). But we already have  $\Omega \cong M(q^k)$ . Hence  $f=1$  and  $\Omega = M(q^k)$  and  $G_\Omega = G_{M(q^k)}$ .

[Case  $A_{q>3}$ .] From (7) for  $\kappa=1$  we have  $\gamma_j \lambda_{j,1} = -4 \neq 0$ ; hence  $\gamma_j$  is a  $\kappa$ ,  $\gamma_j = \kappa_j$ . Unless  $q^k=3^1$ ,  $\kappa$  has a value  $\kappa \neq \pm 1$ , and for such a  $\kappa$  we see from (7) that  $\alpha_j$  is a  $\lambda$ ,  $\alpha_j = \lambda_j$ . Also in case  $q^k=3^1$  we have  $\alpha_j = \lambda_j$ . Hence in every case  $\delta_j = 2 - \alpha_j$  is a  $\lambda$  and since  $\alpha_j \delta_j - \beta_j \gamma_j = 1$  so is also  $\beta_j$  a  $\lambda$ .

[Case  $A_{q=2}^{k \geq 1}$ .]  $\kappa$  has a value  $\kappa \neq 1$  ( $+1 = -1$ ) and for such a  $\kappa$  we see from (7) <sub>$q=2$</sub>  that  $\alpha_j / \gamma_j$  is a  $\lambda$ , say  $\alpha_j / \gamma_j = \lambda_j$ . Since  $q=2$ ,  $\delta_j = \alpha_j$  and  $\delta_j / \gamma_j = \lambda_j$ . The substitutions  $V = V_j V_{\kappa, \lambda}$  with  $j$  fixed have  $\alpha_j / \gamma_j = \lambda_j$ . If two substitutions  $V_{(1)} = V_{j_1} V_{\kappa, \lambda_1}$  and  $V_{(2)} = V_{j_2} V_{\kappa, \lambda_2}$  have  $\lambda_{j_1} = \lambda_{j_2}$ , then  $j_1 = j_2$ . There are  $f q^k$  substitutions  $V_j$ ;  $f q^k > 2$ . Setting  $V_{j_1} V_{j_2} = V_{j_3} V_{\kappa, \lambda}$ , where  $j_1 \neq j_2$  and so  $\lambda_{j_1} \neq \lambda_{j_2}$ ,  $\lambda_{j_1} + \lambda_{j_2} \neq 0$ , we find

$$\lambda_{j_3} = \frac{\lambda_{j_1} \lambda_{j_2}}{\lambda_{j_1} + \lambda_{j_2}} + \frac{1}{\lambda_{j_1} + \lambda_{j_2}} \frac{\beta_{j_1}}{\gamma_{j_1}}.$$

Hence every  $\beta_j / \gamma_j$  is a  $\lambda$ ,  $\beta_j / \gamma_j = \lambda'_j$ . Since  $\alpha_j \delta_j - \beta_j \gamma_j = 1$ ,  $\gamma_j^2 (\lambda_j^2 - \lambda'_j) = 1$ , and so  $\gamma_j^2$  is a  $\lambda$  and hence (since  $q=2$ ) so is  $\gamma_j$  a  $\lambda$  and with  $\gamma_j$  also  $\alpha_j, \delta_j, \beta_j$ .

**Theorem:** The group  $G_\Omega$  ( $q^k=2^1$ ) is a dihedral group  $G_\Omega$  whose order is  $\Omega = 2(1+2f)$  (§ 9).

[Case  $A_{q=2}^{k \geq 1}$ .] The group  $G_\Omega$  ( $q^k=2^1$ ) of order  $2(1+2f)$  is obtained by extending  $G_{2^1}^{(2)} \{S_\lambda\}$ , where  $\lambda = [1] = 0, 1$ , by certain  $2f$  substitutions  $V_j$  of period 2 ( $j=1, \dots, 2f$ ). The  $1+2f$  substitutions  $V_j$  ( $j=0, \dots, 2f$ ) ( $V_0=S_1$ ) form one conjugate set under the  $G_\Omega$ . Setting  $U_j = V_j V_0$ , so that  $V_j = U_j V_0$ , we have  $I = V_j^2 = U_j V_0 U_j V_0 = U_j V_0 U_j V_0^{-1}$ , so that  $V_0 U_j V_0^{-1} = U_j^{-1}$  ( $j=0, \dots, 2f$ ). A substitution  $U$  transforms under  $G_\Omega$  to a  $U$ . Hence for  $V_0$  we may substitute  $V_j$  and have  $I = U_j V_j U_j V_j$ . Hence  $U_j V_j = U_j U_j V_0$  is of period 2 and so is a  $V_j = U_j V_0$ . Hence  $U_j U_j = U_j$ . The substitutions  $U_j$  form a group  $G_{1\Omega}$ . We have  $U_j U_{j'} = (U_j^{-1} U_{j'}^{-1})^{-1} = (U_{j''})^{-1} = V_0 (U_{j''}) V_0^{-1} = V_0 (U_{j'}^{-1} U_j^{-1}) V_0^{-1} = (V_0 U_j^{-1} V_0^{-1}) (V_0 U_{j'}^{-1} V_0^{-1}) = U_j U_{j'}$ . The group  $G_{1\Omega} = G_{1+2f}$  is commutative and of odd order. Hence in our case (§ 6 II, III) it is cyclic. Since  $V_0 U_j V_0^{-1} = U_j^{-1}$ , the group  $G_\Omega$  is a dihedral  $G_{2(1+2f)}$  based on a cyclic  $G_{1+2f}$  of set II or III. This dihedral group  $G_\Omega$  was exhibited in § 9. We obtain as a particular case (for  $f=1$ ) of the dihedral  $G_{2(1+2f)}$  the  $G_{M(2)}$  called for by the preceding theorem.

(5°) [Case B.]  $q > 2$ ,  $n/k = \text{even}$ .  $\Omega = (1+f q^k) q^k (q^k-1)$ .  $q^k-1 = d_-$ .

\* Case A:  $q^k=3^1$ ,  $d_- = 1$ ,  $\lambda=0, +1, -1$ .  $\kappa = \eta = \gamma_j = 1, -1$ . We may and do restrict  $\eta$  and the  $\gamma_j$  to the value 1.

$V_j V_{\eta, \lambda} = V_j S_\lambda$  has as its  $\alpha + \delta$  the mark  $\alpha_j + \delta_j + \lambda$ . For  $\lambda=0, +1, -1$  two of the three values  $\alpha_j + \delta_j + \lambda$  are  $+2, -2$ . Hence  $\alpha_j + \delta_j + \lambda$  has the values  $0, +1, -1$ . By convenient initial choice of the  $V_j$  we have  $\alpha_j + \delta_j = 0$ . The general extender  $V_j = V_{(a)}$  is then

$$V_{(a)} = \begin{pmatrix} a & -1-a^2 \\ 1 & -a \end{pmatrix}.$$

Every substitution of the group  $G_\Omega = G_{2(1+2f)}$  has its  $\gamma$  a  $\lambda$  and its  $\alpha + \delta$  a  $\lambda$ . Further, by transformation by  $S_{-a_0}$  (cf. (2°) above), where  $a_0$  is any particular  $a$ , since  $S_{-a_0} V_{(a)} S_{-a_0}^{-1} = V_{(a-a_0)}$ , we arrange it so that one  $a$  is  $a=0$ .

$$V_{(0)} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad V_{(a)} = \begin{pmatrix} -1 & -a \\ a & -1-a^2 \end{pmatrix}.$$

Hence every  $a$  and the  $\gamma$  of  $V_{(0)} V_{(a)}$  is a  $\lambda$ .

The  $2d_-$  marks  $\eta$  are the square roots of the  $q^k-1$  marks  $\kappa$ , that is, they are the distinct powers of  $\eta_0 = \sqrt{\kappa_0}$ , where  $\kappa_0$  is a primitive root of the Galois-field  $GF[q^k]$ . In particular there is a mark  $\eta = \sqrt{-1}$ .

We analyze the  $G_\Omega$  of this case  $B$  with respect to its substitutions  $T$  of period 2 (of set II).

The  $G_\Omega$  contains the  $G_{q^k(q^k-1)}^{(\infty)}$  as one of a system of  $1 + fq^k$  conjugate groups; in  $G_{q^k(q^k-1)}^{(\infty)}$  there are  $q^k$  conjugate cyclic  $G_{q^k-1=d_-}^{(\infty)}$ , and so in all  $q^k$  substitutions  $T$ , and these are all conjugate with

$$T_0 = (\sqrt{-1}, 0; 0, -\sqrt{-1}) . \quad (9)$$

Under the  $G_\Omega$   $T_0$  is one of a system of  $(1 + fq^k) q^k$  or  $\frac{1}{2} (1 + fq^k) q^k$  conjugate substitutions  $T$  according as  $T_0$  is within  $G_\Omega$  self-conjugate under the cyclic  $G_{q^k-1}^{(\infty)}$  or under a dihedron  $G_{2(q^k-1)}$  obtained by extending it by a substitution  $T'_0$  interchanging  $\infty$  and  $0$  (§ 6 II, § 9). There are then amongst the substitutions  $V_j V_{\eta, \lambda}$  ( $j > 0$ ) at least  $f q^{2k}$  or  $\frac{1}{2} (f q^{2k} - 1) q^k$  substitutions  $T = V_j V_{\eta, \lambda}$  satisfying the necessary and sufficient condition (§ 6 end)

$$\alpha_j \eta + \delta_j \eta^{-1} + \gamma_j \eta \lambda = 0 . \quad (10)$$

But there are at most  $f q^k (q^k - 1)$  such substitutions  $V_j V_{\eta, \lambda}$ . Hence the latter alternative holds; the  $G_\Omega$  contains a substitution

$$T'_0 = (0, -\tau^{-1}; \tau, 0) , \quad (11)$$

and  $f$  is odd (since  $\frac{1}{2} (f q^{2k} - 1) q^k$  is an integer).

In case a  $V_j$  ( $j > 0$ ) gives rise to one or more substitutions  $T = V_j V_{\eta, \lambda}$  we replace it by one such  $T$ . For these  $V_j = T_j$  then we set  $\delta_j = -\alpha_j$ , and further  $\gamma_j = 1$ . We attend for the moment only to these  $V_j$ 's. The substitutions  $V_j V_{\eta, \lambda}$  derived from such  $V_j$  have  $\alpha/\gamma = \alpha_j/\gamma_j$  as a characteristic invariant. The equation  $T = V_j V_{\eta, \lambda}$  has always one solution:  $V_{\eta, \lambda} = I$ ,  $T = V_j$ . If it has two solutions, then the invariant  $\alpha_j/\gamma_j$  is a  $\lambda$ ,  $\alpha_j/\gamma_j = \lambda_j$ . Conversely, whenever  $\alpha_j/\gamma_j = \lambda_j$ , then the equation has exactly  $q^k - 1$  solutions.

By comparing the minimum and maximum numbers for the  $T = V_j V_{\eta, \lambda}$  ( $j > 0$ ) in the  $G_\Omega$  we have

$$\frac{1}{2} (f q^{2k} - 1) q^k \leq q^k (q^k - 1) + (f - 1) q^k . \quad (12)$$

There are in all two cases,  $f = 1$ ,  $f = 3$ :

$$[B_{f=1}] \quad q > 2, n/k = \text{even}, \Omega = (q^k + 1) q^k (q^k - 1) = 2M(q^k).$$

$$[B_{f=3}] \quad q = 3, k = 1, n = \text{even}, \Omega = 60.$$

[Case  $B_{f=1}$ ] *Theorem: The group  $G_\Omega$  is the group<sup>n</sup>  $G_{2M(q^k)}$  of  $\Omega = 2M(q^k) = (q^k + 1) q^k (q^k - 1)$  substitutions*

$$V = (\lambda_1, \lambda_2; \lambda_3, \lambda_4) \quad (\lambda_1 \lambda_3 - \lambda_2 \lambda_4 \neq 0) \quad (13)$$

where the  $\lambda_1, \lambda_2, \lambda_3, \lambda_4$  belong to the  $GF[q^k]$  of which  $\eta_0^2 = \kappa_0$  is a primitive root. The

<sup>n</sup> For  $q = 3$  the group  $G_{2M(q^k)}$  is of octahedron type (§ 1).

group  $G_{M(q^k)}$  is obtained by extending its self-conjugate subgroup  $G_{M(q^k)} \{V\}$  of substitutions  $V$  of square determinant by  $P_{\eta_0}$ ,

$$P_{\eta_0} = (\eta_0, 0; 0, \eta_0^{-1}) = (\kappa_0, 0; 0, 1) = D_{\kappa_0}. \quad (14)$$

We have in the  $T_\lambda$ ,

$$T'_\lambda = S_\lambda T'_0 S_{-\lambda} = \begin{pmatrix} 1, \lambda \\ 0, 1 \end{pmatrix} \begin{pmatrix} 0, -\tau^{-1} \\ \tau, 0 \end{pmatrix} \begin{pmatrix} 1, -\lambda \\ 0, 1 \end{pmatrix} = \begin{pmatrix} \lambda \tau, -\lambda^2 \tau - \tau^{-1} \\ \tau, -\lambda \tau \end{pmatrix}, \quad (15)$$

$q^k = f q^k$  substitutions  $T = V_j V_{\eta, \lambda}$  of distinct invariants  $a/\gamma = \lambda \tau / \tau = \lambda$ . We choose these as the  $q^k = f q^k$  extenders  $V_j$  and set  $a_j/\gamma_j = \lambda_j$ . We find as in 4° [Case  $A_{q=2}^{k \geq 1}$ ] that  $\beta_j/\gamma_j = \lambda'$  and  $\gamma'_j = \lambda'_j = \kappa_j$ ; whence  $\gamma_j = \eta_j$ . Thus  $V_j$  has the form

$$V_j = (\lambda_j \eta_j, \lambda'_j \eta_j; \eta_j, \lambda''_j \eta_j). \quad (16)$$

$V_j$  or  $V_j V_{\eta_0, 0}$  has the form (13) according as  $\eta_j$  is an even or an odd power of  $\eta_0 = \sqrt{\kappa_0}$ ; we denote it by  $V'_j$ .

When we extend the  $G_{M(q^{k-1})}^{(\kappa)} \{V_{\kappa, \lambda}\}$  by these  $q^k V'_j$  we obtain the  $G_{M(q^k)}$  of the caption theorem. Since  $P_{\eta_0}^2 = P_{\eta_0} = P_{\kappa_0}$  and  $P_{\eta_0} G_{M(q^k)} P_{\eta_0}^{-1} = G_{M(q^k)}$ , this group  $G_{M(q^k)}$  extends by  $P_{\eta_0} = V_{\eta_0, 0}$  to the  $G_{M(q^k)}$  of the caption theorem which is therefore our group  $G_0$  of order  $\Omega = 2M(q^k)$ .

[Case  $B_{f=3}$ ] *Theorem.* The group  $G_{\Omega=60}$  is a subgroup of the group  $G_{M(3)} = G_{360}$  of the type  $A(8)$  depending on the  $GF[3^2]$ . It is generated by the generators

$$E_1 = S_{-1} = \begin{pmatrix} 1, -1 \\ 0, 1 \end{pmatrix}, \quad E_2 = T_{0, \sqrt{-1}} = \begin{pmatrix} \sqrt{-1}, 0 \\ 1, -\sqrt{-1} \end{pmatrix}, \quad E_3 = P_{\sqrt{-1}} = \begin{pmatrix} \sqrt{-1}, 0 \\ 0, -\sqrt{-1} \end{pmatrix} \quad (17)$$

subject to the generational relations

$$I = E_1^3 = E_2^3 = E_3^3 = (E_1 E_2)^3 = (E_1 E_3)^3 = (E_2 E_3)^3, \quad (18)$$

and hence (§ 1) is an icosahedron group.

We have  $q = 3, k = 1, n = \text{even}, \Omega = 60, f = 3$ . The  $q^k - 1 = 2$  marks  $\kappa$  are  $\kappa = +1, -1$ ; the marks  $\lambda$  are  $\lambda = 0, +1, -1$ ; the marks  $\eta$  are  $\eta = +1, -1, +\sqrt{-1}, -\sqrt{-1}$ .  $\kappa_0 = -1$ ;  $\eta_0 = \sqrt{-1}$ .

The  $G_0$  contains the group  $G_{M(q^{k-1})}^{(\kappa)} = G_0^{(\kappa)} \{V_{\kappa, \lambda}\}$ , which is obtained by extending the  $G_0^{(\kappa)} \{V_{\kappa, \lambda} = S_\lambda\}$  by  $T_0 = P_{\eta_0}(9)$ , and also the substitution  $T'_0$  (11). In  $T'_0$   $\tau$  is not a mark  $\eta$ ; if it were, the  $G_{\Omega=60}$  would contain (§ 7) a subgroup  $G_{2M(3)} = G_{24}$  (13).

The  $G_0^{(\kappa)} \{V_{\kappa, \lambda}\}$  extends to the  $G_{\Omega=60}$  by the identity and certain nine extenders  $V_j$ ; since in (12) the equality holds we may take every extender as a substitution  $T = T_j$  of period 2; then the three extenders  $T_j = T'_j = T'_\lambda$  (15) ( $j = 1, 2, 3$ ;  $\lambda = 1, -1, 0$ ) with invariants  $\lambda = 1, -1, 0$  yield each two substitutions  $T = T_j V_{\eta, \lambda}$ , while the remaining six extenders  $T_j$  ( $j = 4, \dots, 9$ ) yield each one substitution  $T = T_j V_{\eta, \lambda}$ .

Analyzing the  $G_{\Omega=60}$  with respect to its 18 substitutions  $V_j V_{\eta, \lambda}$  of set I (3°), and remembering that an extender  $V_j$  yields at most  $2d_- = 4$  such substitutions  $V_j V_{\eta, \lambda}$  (3°), we find in our  $G_0$  that (1) no extender  $V_j = T_j$  yields 4, (2) the three extenders

$T'_\lambda$  yield 0, and hence (3) the remaining six extenders  $T_i$  yield each 3 such substitutions of set I, and then (4) these six  $T_i$  are exactly the six  $T = T_{\lambda, \eta'}$ , with  $(\alpha, \beta, \gamma, \delta) = (\lambda + \eta', \beta, 1, -\lambda - \eta')$  where  $\alpha\delta - \beta\gamma = 1$  for  $\lambda = 0, +1, -1$ ;  $\eta' = \eta_0, -\eta_0$ .

The  $G_{\Omega=60}$  is then, if indeed it exist at all, determined uniquely by the  $G_s^{(n)} \{V_{\eta, \lambda}\}$  and the six extenders  $T_{\lambda, \eta'}$  as a subgroup of the  $G_{M(n)=200}$  of type A (8) depending upon the  $GF[3^3]$ ; this group is a subgroup of the fundamental  $G_{M(n)}$ , since  $n$  is even.

Now the three substitutions (17) which satisfy the relations (18) generate (by § 1) an icosahedron group  $G_{60}$ . This actually existing icosahedron subgroup  $G_{60}$  of the  $G_{M(n)}$  ( $n = \text{even}$ ) is easily seen to be the (hypothetical)  $G_{\Omega=60}$  of our present analysis [Case B,  $f = 3$ ].

(6°) *The systems of conjugate groups determined by the various  $G_\Omega(2^\circ, \dots, 5^\circ)$ .*

We have  $(2^\circ, \dots, 5^\circ)$  separated the subgroups  $G_\Omega$  of the  $G_{M(n)}$  of linear fractional substitutions of determinant 1 in the  $GF[s = q^n]$  (§ 4) which contain substitutions of period  $q$  and are subgroups of no one of the  $s + 1$   $G_{\frac{s(s-1)}{2; 1}}^{(\mu)}$  into four classes I  $\dots$  IV.

The subgroups  $G_{\Omega I} \dots G_{\Omega IV}$  of these respective classes are under the  $G_{(\alpha; 1) M(n)}$  of all substitutions in the  $GF[s]$  (§ 4) conjugate with

- (I) [A] a  $G_{I(k)} = G_{M(q^k)}$  (8), where  $k$  is a divisor of  $n$ ;
- (II) [ $B_{f=1, q>2, n/k=\text{even}}$ ] a  $G_{II(k)} = G_{M(q^k)}$  (13), where  $k$  is a divisor of  $n$  with even  $n/k$ ;
- (III) [ $A_{q^k=2^l}$ ] a dihedron  $G_{III(f)} = G_{2(1+2f)}$ , where  $1 + 2f$  is a divisor of  $q^n - 1 = 2^n - 1$ ;
- (IV) [ $B_{f=2, q=2, k=1, n=\text{even}}$ ] the  $G_{IV} =$  the icosahedron  $G_{60}$  (17).

These typical groups  $G_{I(k)}, G_{II(k)}, G_{III(f)}, G_{IV}$  actually exist under the conditions just now specified.

The  $G_{I(k)} = G_{M(q^k)}$  is simple (§ 8).

The  $G_{II(k)} = G_{2M(q^k)}$  contains  $G_{M(q^k)} = G_{I(k)}$  as the only proper self-conjugate subgroup for  $q^k > 3$ ; the  $G_{2M(q^k)}$  is an octahedron group  $G_{24}$  (§ 1), and contains the self-conjugate  $G_{M(q^k)}$  and also a self-conjugate four-group.

The  $G_{III(f)} = G_{2(1+2f)}$  contains as its self-conjugate subgroups the cyclic subgroups of its cyclic base  $G_{1+2f}$ .

The  $G_{IV} = G_{60}$  is simple.

Within the  $G_{M(n)}$  any such group  $G_\Omega$  is self-conjugate under a  $G_\Omega$ , which is likewise a group  $G_\Omega$ . The two groups  $G_\Omega, G_\Omega$  may be simultaneously transformed each to its typical form (I  $\dots$  IV).

Within the  $G_{M(n)}$  each group  $G_{\Omega II}, G_{\Omega III}, G_{\Omega IV}$  is self-conjugate under itself and so is one of a system of  $M(s)/\Omega$  conjugate groups.

Since the  $G_{I(k)}$  is self-conjugate under the  $G_{II(k)}, G_{I(k)}; G_{I(k)}$  according as  $(q > 2)$   $n/k = \text{even}, n/k = \text{odd}; (q = 2)$ , each group  $G_{\Omega I(k)}$  is one of a system of  $M(s)/(2, 1; 1)\Omega$  conjugate groups.

The question remains as to the number of such conjugate systems in the various cases, for we have (2°) introduced a reducing transformer  $P_{\sqrt{\lambda_0}}$  which is in the  $G_{(2;1)M(n)}$ , and in the  $G_{M(n)}$  only if the mark  $\lambda_0$  is a square in the  $GF[q^n]$ . There is then only one such system of conjugate groups  $G_n$  if  $q = 2$  ( $G_{nI}$ ,  $G_{nIII}$ ) (since then every mark  $\lambda_0$  ( $\lambda_0 \neq 0$ ) is a square) or if  $q > 2$ ,  $n/k = \text{odd}$  ( $G_{nI}$ ) (since then every  $[\lambda_1, \dots, \lambda_{m=k}]$  with the  $GF[q^k]$  as multiplier field has besides the mark 0 half its marks squares and half not-squares, and since we may choose for  $\lambda_0$  any one of its square marks). But there are two such systems of conjugate groups  $G_n$  ( $G_{nI}$ ,  $G_{nII}$ ,  $G_{nIV}$ ) if  $q > 2$ ,  $n/k = \text{even}$ , which interchange under the transformer  $P_{\sqrt{\nu}} = D$ , where  $\nu$  is any not-square in the  $GF[q^n]$ .

The results for  $q > 2$  just stated may well be obtained otherwise by consideration within the  $G_{M(n)}^{s+1}$ , of which the  $G_{M(n)}^{s+1}$  and the  $G_{2M(n)}^{s+1}$  are subgroups. If  $q > 2$ ,  $n/k = \text{even}$ , then each group  $G_{nI(k)}$ ,  $G_{nII(k)}$ ,  $G_{nIV}$  is within the  $G_{M(q^n)}$ , or within the  $G_{M(q^{2n})}$ , self-conjugate under the same group; hence every system of conjugate groups under the  $G_{2M(n)}$  breaks into two systems of conjugate groups under the  $G_{M(n)}$ . But if  $q > 2$ ,  $n/k = \text{odd}$ , and so  $q > 2$ ,  $2n/k = \text{even}$ , then each group  $G_{nI(k)} = G_{M(q^k)}$  is within the  $G_{M(q^n)}$ , self-conjugate under itself but in the  $G_{M(q^{2n})}$ , self-conjugate under a  $G_{2M(q^k)} = G_{nII(k)}$ ; hence every system of conjugate groups under the  $G_{2M(n)}$  remains a system of conjugate groups under the  $G_{M(n)}$ .

(7°) *The subgroups  $G_n$  (2° . . . 5°) of the  $G_{2M(n)}$  ( $q > 2$ ) and their systems of conjugate groups.*

We find that the  $G_{2M(n)}$  contains the representative groups

- (I)  $[A]$  a  $G_{I(k)} = G_{M(q^k)}$  (8), where  $k$  is a divisor of  $n$ ;
- (II)  $[B_{f=1, q>2}]$  a  $G_{II(k)} = G_{2M(q^k)}$  (13), where  $k$  is a divisor of  $n$ ;
- (IV)  $[B_{f=2, q=2, k=1, n=\text{even}}]$  the  $G_{IV} = G_{\infty}$  (17),

whose systems contain respectively (I)  $2M(s)/2\Omega I(k)$ ; (II)  $2M(s)/\Omega II(k)$ ; (IV)  $2M(s)/\Omega IV$  groups, every  $G_{I(k)}$  being self-conjugate under a corresponding  $G_{II(k)}$ .

(8°) *The tetrahedron and the octahedron subgroups  $G_n$  of the  $G_{M(n)}$ .*

The only typical subgroups  $G_n$  (I . . . IV) of order 12 and 24 are (I) the  $G_{M(n)}^{s+1} = 12$  of the  $G_{M(n)}$ , and (II) the  $G_{2M(n)}^{s+1} = 24$  of the  $G_{M(n)}$ , ( $n = \text{even}$ ). These are respectively a tetrahedron and an octahedron group (§1). Each contains a single self-conjugate four-group. The groups  $G_{12}$ ;  $G_{24}$  of these types are for  $n = \text{odd}$ ; even the largest groups containing four-groups as self-conjugate subgroups (§9 end). The statements (of 6°) as to systems of conjugate groups  $G_{12}$ ,  $G_{24}$  for  $q = 3$  are given below (§15, 1°).

(9°) *The icosahedron subgroups  $G_n$  of the  $G_{M(n)}$ .*

The only typical subgroups  $G_n$  (I . . . IV) of order 60 are (I) the  $G_{M(n)}^{s+1} = 60$  of the  $G_{M(n)}$ , (I) the  $G_{M(n)}^{s+1} = 60$  of the  $G_{M(n)}$ , ( $n = \text{even}$ ), and (IV) the  $G_{IV}$  of the  $G_{M(n)}$ , ( $n = \text{even}$ ). Each of these is an icosahedron group (§1). The statements (of 6°) as to systems of conjugate groups  $G_{60}$  are given below (§16, 3°).

§14. SUBGROUPS OF  $G_{M(q)}$  CONTAINING NO OPERATORS OF PERIOD  $q$ (1°) *The diophantine equations.*<sup>28</sup>

The  $G_\Omega$  contains only substitutions of sets II and III (§ 6). Every substitution of  $G_\Omega$  lies in and determines within  $G_\Omega$  a largest cyclic subgroup  $G_d$ . Two such groups  $G_d$  have only the identity in common. Every  $G_d$  is within  $G_\Omega$  one of a system of  $\Omega/fd$  conjugate groups, where  $f$  is either 1 or 2, according as the  $G_d$  is self-conjugate within itself or within a dihedron  $G_{2d}$  based on itself.

The  $G_\Omega$  contains say  $r$  such systems. Then the enumeration of the  $\Omega$  substitutions of  $G_\Omega$  leads to the relations

$$\Omega = 1 + \sum_{i=1}^r (d_i - 1) \Omega / f_i d_i \quad (f_i = 1 \text{ or } 2), \quad (1)$$

$$\Omega \geq f_i d_i \quad (i = 1, 2, \dots, r). \quad (2)$$

Further, if two non-conjugate cyclic groups  $G_{d_i}, G_{d_j}$  of odd order are present, then there are at least  $d_j$  groups in the system determined by  $G_{d_i}$ , and *vice versa*, so that

$$\Omega \geq 2 d_i d_j - d_i - d_j + 1. \quad (3)$$

The discussion of the relations (1, 2, 3) leads to the six cases:

Case	$r$	$d_1, d_2, d_3$	$f_1, f_2, f_3$	$\Omega$
I.....	1	$d$	1	$d$
II.....	2	$d, 2$	2, 1	$2d$
III.....	3	$d, 2, 2$	2, 2, 2	$2d$
IV.....	2	3, 2	1, 2	12
V.....	3	4, 3, 2	2, 2, 2	24
VI.....	3	5, 3, 2	2, 2, 2	60

(2°) We discuss these cases more narrowly from our present group-theoretic standpoint.

[Case I.] The  $G_{\Omega=d}$  is a cyclic  $G_d$  of § 9.

[Case II.] The  $G_{\Omega=2d}$  is a dihedron  $G_{2d}$  with  $d$  odd of § 9.

[Case III.] The  $G_{\Omega=2d}$  is a dihedron  $G_{2d}$  with  $d$  even of § 9.

[Case IV.] This case occurs only for  $q \neq 2, 3$ . The  $G_{\Omega=12}$  contains a system of four groups  $G_3$ , each self-conjugate exactly under itself. These contain only the identity in common. Accordingly the  $G_{12}$  is holodrically isomorphic with a substitution-group  $G_{12}^4$  of order 12 on 4 letters. This  $G_{12}^4$  is the alternating  $G_{12}^4$ , and so the  $G_{12}$  is a *tetrahedron group*. It contains one self-conjugate four-group  $G_4$ —the dihedron  $G_4$  in which each of the three conjugate  $G_{d_2=2}$  is self-conjugate.

[Case V.] This case occurs only for  $q \neq 2, 3$ . The  $G_{\Omega=24}$  contains a system of four cyclic groups  $G_3$  and correspondingly a system of four dihedron groups  $G_4$ . These

<sup>28</sup>The equation (1) originated in C. JORDAN's analytic attack (*Crelle*, Vol. LXXXIV, 1878) on the problem of the determination of all finite groups of binary and ter-

nary linear substitutions. Cf. GIERSTER, *loc. cit.*, pp. 358 f., and KLEIN-FRICKER, *loc. cit.*, pp. 483 f.

$G_4$  contain only the identity in common.<sup>29</sup> Accordingly the  $G_{24}$  is holoedrically isomorphic with a substitution-group  $G_{24}^4$  of order 24 on 4 letters. This  $G_{24}^4$  is the symmetric group  $G_{24}^4$ , and so the  $G_{24}$  is an octahedron group. It contains one self-conjugate tetrahedron group  $G_{12}$  and one self-conjugate four-group  $G_4$ . (This  $G_4$  contains the identity, and the three substitutions  $V_4', V_4'', V_4'''$  of period two, where  $V_4', V_4'', V_4'''$  are generators of the three conjugate cyclic  $G_4$ . For  $V_4', V_4'', V_4'''$  form a system of conjugate substitutions under the  $G_{24}$ ; and they are commutative, since  $V_4'$  transforms  $G_4\{V_4''\}$  into  $G_4\{V_4'''\}$  and so  $V_4'^2$  transforms  $G_4\{V_4''\}$  into itself and hence  $V_4''^2$  into itself.)

[Case VI.] This case occurs only for  $q \neq 2, 3, 5$ . The  $G_{60}$  contains a system of 15 conjugate substitutions of period 2. Each is self-conjugate in a four-group  $G_4$ . There are 5 such four-groups; they are conjugate under the  $G_{60}$ . Each  $G_4$  is self-conjugate in a  $G_{12}$  of tetrahedron<sup>30</sup>-type (Case IV). There are five conjugate  $G_{12}$ , since each contains only one four-group. These  $G_{12}$  have only the identity in common. Hence the  $G_{60}$  is holoedrically isomorphic with a substitution-group  $G_{60}^5$  of order 60 on 5 letters. This  $G_{60}^5$  is the alternating  $G_{120}^5$ , and so the  $G_{60}$  is an icosahedron group.

#### § 15. THE TETRAHEDRON AND THE OCTAHEDRON SUBGROUPS OF THE $G_{M(n)}$

(1°) From § 13 (8°) and § 14 (Cases III and IV) and the four-group considerations of § 9 and for  $q = 2$  from § 12 (2°) we see easily:

$[s = q^n = 8h \pm 1.]$  The  $G_{M(n)}$  contains two systems of  $M(s)/24$  conjugate octahedron groups  $G_{24}$  and two systems of  $M(s)/24$  conjugate tetrahedron groups  $G_{12}$ . Every  $G_{12}$  is self-conjugate under a  $G_{24}$ . The two systems coalesce into one system within the  $G_{2M(n)}$ . The  $G_{2M(n)}$  contains no other tetrahedron or octahedron groups.

$[s = q^n = 8h \pm 3]$  and  $[s = 2^n (n \text{ even}).]$  The  $G_{M(n)}$  contains no octahedron group  $G_{24}$  and one system of  $M(s)/12$  conjugate tetrahedron groups  $G_{12}$ . For  $q > 2$  the  $G_{2M(n)}$  contains one system of  $M(s)/12$  conjugate octahedron groups  $G_{24}$  containing respectively these  $G_{12}$ . The  $G_{2M(n)}$  contains no other tetrahedron or octahedron groups.

For  $q > 2$  the  $G_{2M(n)}$  always contains tetrahedron  $G_{12}$  and octahedron  $G_{24}$ , and the  $G_{M(n)}$  always contains the  $G_{12}$ , and, when 24 is a divisor of  $M(s)$ , also the  $G_{24}$ .

For  $q = 2$ , although the order  $M(2^n)$  is a multiple of 12 for  $n > 1$  and of 24 for  $n > 2$ , the group  $G_{M(2^n)}$  contains tetrahedron  $G_{12}$  only for  $n$  even, and octahedron  $G_{24}$  never.

(2°) We exhibit generators of representatives of these systems of groups conjugate within the  $G_{(2;1)M(n)}$ .

$[s = 2^n, n \text{ even}]$   $G_{12} \{E_1, E_2\}$  of the  $G_{M(2^n)}^{2^n+1}$ :

$$E_1 = (\kappa, 0; 0, \kappa^{-1}), \quad E_2 = (1, 1; 0, 1) \quad (\kappa^2 + \kappa + 1 = 0).$$

<sup>29</sup>The four  $G_4$  contain only the identity in common. The four  $G_4$  do not contain two substitutions of period 2 in common, since then the four  $G_4$  and so the four  $G_{24}$  would be identical. The four  $G_4$  do not contain in common one single substitution of period 2, since this substitution would be

self-conjugate in the  $G_{24}$  and this  $G_{24}$  contains no self-conjugate substitution.

<sup>30</sup>Since no  $G_{12}$  of either of cases VI, V, III, II, I contains a self-conjugate four-group.



$[q > 2]$   $G_{12} \{E_1, E_2\}$ ,  $G_{24} \{B_0, B_1, B_2\}$  of the  $G_{12}^{s+1}$  or of the  $G_{24}^{s+1}$ , according as  $s = q^n = 4k + 1$  or  $4k - 1$ :

$$E_1 = \begin{pmatrix} \frac{1+\sqrt{-1}}{2} & \frac{-(1+\sqrt{-1})}{2} \\ \frac{1-\sqrt{-1}}{2} & \frac{1-1-1}{2} \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

$$B_0 = (1, 1; 1, -1) = (\sqrt{\rho}, \sqrt{\rho}; \sqrt{\rho}, -\sqrt{\rho}), \quad B_1 = B_0 E_1, \quad B_2 = B_0 E_2.$$

The case  $q = 2$  is given in § 12, 3°.

The case  $q > 2$  is treated thus: Attending to the properties of the abstract tetrahedron and octahedron groups (§ 1) and starting from the representative four-group (§ 9) of the  $G_{M(u)}$ —in suitable form—

$I, K = (\sqrt{-1}, 0; 0, -\sqrt{-1}), L = (0, -1; 1, 0), M = KL = (0, \sqrt{-1}; \sqrt{-1}, 0)$ , we set down for an operator  $U$  of the  $G_{M(u)}$  the system of equations

$$U^2 = I, \quad UK = LU, \quad UL = MU, \quad UM = KU,$$

and find the four solutions

$$U = U_{e_1, e_2} = \begin{pmatrix} \frac{1+e_1\sqrt{-1}}{2} & \frac{-e_2(1+e_1\sqrt{-1})}{2} \\ \frac{e_2(1-e_1\sqrt{-1})}{2} & \frac{1-e_1-1}{2} \end{pmatrix} \quad (e_1 = \pm 1, e_2 = \pm 1),$$

which all belong to the tetrahedron group

$$G_{12} \{E_1, E_2\} : E_1 = U_{1,1}, \quad E_2 = L : I = E_1^2 = E_2^2 = (E_1 E_2)^2.$$

Then the substitution  $B_0$  of the  $G_{24 M(u)}$  uniquely determined by the equations

$$I = B_0^2 = (B_0 E_1)^2 = (B_0 E_2)^2$$

extends the  $G_{12} \{E_1, E_2\}$  to the octahedron group

$$G_{24} \{B_0, B_1, B_2\} : I = B_0^2 = B_1^2 = B_2^2 = (B_0 B_1)^2 = (B_0 B_2)^2 = (B_1 B_2)^2.$$

The  $G_{24} \{B_0, B_1, B_2\}$  does or does not belong to the  $G_{M(u)}$  according as  $-1, -2$  are or are not of the same quadratic character (square or not-square) in the  $GF[q^n]$ . The supplementary quadratic reciprocity theorem

$$\left(\frac{2}{q}\right) = (-1)^{\frac{q^2-1}{4}}$$

exhibits the accord of this statement with that of (1°).

In case within the  $G_{M(u)}$  there are *two* systems of conjugate groups a representative of the second system is obtained from a representative of the first by transforming say by  $D_\rho$ .

#### § 16. THE ICOSAHEDRON SUBGROUPS OF THE $G_{M(u)}$

(1°) We seek within the  $G_{M(u)}$  a system of three generators  $E_1, E_2, E_3$  of an icosahedron group  $G_{60}$  which satisfy the generational relations of § 1

$$I = E_1^5 = E_2^5 = E_3^5 = (E_1 E_2)^3 = (E_1 E_3)^3 = (E_2 E_3)^3. \quad (1)$$

The treatment is such as not to include the case  $q = 3$ . The icosahedron groups of the  $G_{M(n=3^n)}$  have however been determined in § 13, 9°.

The generators  $E_1, E_2$  determine a dihedron group  $G_2$ . Under the  $G_{(3,1)M(n)}$  ( $q \neq 3$ ) they are (§§ 9, 6) respectively simultaneously conjugate with

$$E_1 = \left( \frac{E, 0}{0, E^{-1} = -E - 1} \right), \quad E_2 = T = \left( \frac{0, -1}{1, 0} \right), \quad (2)$$

where we are working in the real  $G_{M(n)}^{s+1}$  or in the second imaginary  $G_{M(n)}^{s's+1}$  (§ 5) according as  $\frac{s-1}{2;1}$  or  $\frac{s+1}{2;1}$  is divisible by 3, or, what must be the same thing,<sup>21</sup> according as the mark  $E$ , ( $E \neq 1$  since  $q \neq 3$ ), satisfying the equations

$$E^2 = 1, \quad E^2 + E + 1 = 0, \quad (2E + 1)^2 = -3, \quad (3)$$

does or does not belong to the  $GF[q^n]$ , i. e., according as  $-3$  is a square or not-square in the  $GF[q^n]$ , if  $q > 3$ , and according as  $n$  is even or odd, if  $q = 2$ .

We seek then a substitution

$$E_2 = (A, B; C, -A) \quad (-AA - BC = 1) \quad (4)$$

of period 2 such that the products

$$E_1 E_2 = \left( \frac{EA, EB}{-(E+1)C, (E+1)A} \right), \quad E_2 E_1 = \left( \frac{B, -A}{-A, -C} \right) \quad (5)$$

shall be each of period 3, and so under the  $G_{M(n)}$  conjugate (§ 6 II, III) with  $E_1$ . The necessary and sufficient conditions (§ 6, IV) are

$$A^2 + BC = -1, \quad (1 + 2E)A = 1, \quad B - C = \pm 1, \quad (6)$$

where the second equation implies that in (4) the proper normalization of  $E_2$  is chosen (§ 4 (7)).

[Case  $q = 2$ .] We find at once two solutions  $E_2$ , viz.:

$$E_2' = (1, 1; 0, 1) \quad E_2'' = (1, 0; 1, 1), \quad (7)$$

and so have two icosahedron groups  $G_{60} \{E_1, E_2', E_3\}$  and  $G_{60} \{E_1, E_2'', E_3\}$ . Since the  $E_2', E_2''$  have neither imaginary form (§ 5 (17, 18)), the groups must be subgroups of the real  $G_{M(n)}$  (if of any  $G_{M(n)}$ ); they are subgroups of the real  $G_{M(n)}$  if and only if  $n$  is even; in that case  $E_1$  is real; indeed the mark  $E$  (3) defines the  $GF[2^2]$  and we determine by either set of three generators the (same) icosahedron  $G_{60} = G_{M(n)}$  of § 13, 8°.

[Case  $q > 3$ .] We find easily

$$A = -\frac{1}{2}(2E + 1), \quad B = \frac{1}{2}(F \pm 1), \quad C = \frac{1}{2}(F \mp 1), \quad A^2 = -\frac{1}{2}, \quad F^2 = -\frac{1}{2}, \quad (8)$$

and have in all four or two sets of generators  $\{E_1, E_2, E_3\}$  according as  $q > 5$  or  $q = 5$ . It is easy to show that these groups  $G_{60} \{E_1, E_2, E_3\}$  are (all) subgroups of the

<sup>21</sup> This (for  $q > 3$ ) is a proof of the reciprocity-formula  $\left(\frac{-3}{q}\right) = \left(\frac{q}{3}\right)$ .

$G_{M(s)}$  in its real or second imaginary form (§5 (17, 18)) if and only if 5 is a square or 0 in the  $GF[q^n]$ , that is, in case  $n$  is even for every prime  $q$ , and in case  $n$  is odd for primes  $q$  of the form  $q = 5$  or  $5k + 1$  or  $5k - 1$ . (We have used the quadratic reciprocity theorem.)

It is important to observe that the  $G_{M(s)}$  contains an icosahedron  $G_{60}$  whenever it contains an operator of period 5, that is, if  $s$  has the form  $5^n$  or  $5k + 1$  or  $5k - 1$ . (The cases  $q = 3$ ,  $n = \text{even}$ , and  $q = 2$ ,  $n = \text{even}$  are included.)

In the  $G_{M(s)}$  ( $q > 3$ ) there are (§9)  $M(s)/6$  dihedral groups  $G_6$  and for each 2.3 generations  $\{E_1, E_3\}$ . In the  $G_{M(s)}$  there are  $M(s)$  dihedral sets  $\{E_1, E_3\}$  and, if any, 4  $M(s)$  or 2  $M(s)$  icosahedron sets  $\{E_1, E_2, E_3\}$  according as  $q > 5$  or  $q = 5$ .

In particular, the icosahedron  $G_{M(s)=60}$  has 60 sets  $\{E_1, E_2\}$  and for each such set 2 sets  $\{E_1, E_2, E_3\}$ . The 60 sets  $\{E_1, E_3\}$  are conjugate under the  $G_{M(s)}$  (§9). The 120 sets  $\{E_1, E_2, E_3\}$  fall into two systems each of 60 sets conjugate under the  $G_{M(s)}$ . Of these two systems the two sets  $\{E_1, E'_2, E_3\}$ ,  $\{E_1, E''_2, E_3\}$

$$E_1 = \left( \frac{E, 0}{0, E^{-1} = -E - 1} \right), \quad E'_2 = \left( \frac{E + 3, \pm 3}{\mp 3, -E - 3} \right), \quad E_3 = \left( \frac{0, -1}{1, 0} \right) \quad (9)$$

are representatives. We notice the relation

$$\left( \frac{0, -1}{1, 0} \right) \left( \frac{A, B}{C, D} \right) \left( \frac{0, 1}{-1, 0} \right) = \left( \frac{-D, C}{B, -A} \right). \quad (10)$$

Hence

$$E_3 E'_2 E_3^{-1} = E'_2. \quad (11)$$

Similar relations hold for the abstract and so for every concrete icosahedron group.

In particular: For  $q = 5$ ,  $n = \text{any}$ , the two sets  $\{E_1, E_2, E_3\}$  (2, 3, 4, 8) yield the same icosahedron group. For  $q > 5$ , the four sets  $\{E_1, E_2, E_3\}$  yield two groups  $G_{60}$  differing only in the choice of  $F$  as root of  $F^2 = -5/3$ . These two groups<sup>22</sup> are conjugate within the  $G_{2M(s)}$  by the transformer  $D_{-1}$  ( $\omega' = -\omega$ ).

Hence, the  $G_{M(s)}$  ( $s = q^n = 5^n$  or  $5k \pm 1$ ,  $q > 3$ ) contains  $M(s)/30$  ( $q > 5$ ), or  $M(s)/60$  ( $q = 5$ ) icosahedron groups  $G_{60}$ , all within the  $G_{2M(s)}$  conjugate with the group  $G_{60} \{E_1, E_2, E_3\}$ ,

$$E_1 = \left( \frac{E, 0}{0, E^{-1} = -E - 1} \right), \quad E_2 = \left( \frac{-\frac{1}{3}(2E + 1), \frac{1}{3}(F + 1)}{\frac{1}{3}(F - 1), \frac{1}{3}(2E + 1)} \right), \quad E_3 = \left( \frac{0, -1}{1, 0} \right), \quad (12)$$

where  $E$  and  $F$  are (any) particular roots of the respective equations.

$$E^2 + E + 1 = 0, \quad 3F^2 + 5 = 0. \quad (13)$$

<sup>22</sup>Since within an icosahedron group  $G_{60}$  the 60 dihedral sets  $\{E_1, E_3\}$  are conjugate, if the dihedral sets  $\{E_1, E_2\}$  of the  $G_{M(s)}$  fall into two systems of conjugate sets (i. e., if  $(s \mp 1)/6$  is even (§9), i. e., if  $s = 12h + 1$  or  $12h + 11$ ), then the icosahedron groups  $G_{60}$  of the  $G_{M(s)}$

must fall correspondingly into two systems, and the two groups  $G_{60} \{E_1, E_2, E_3\}$  containing the same dihedral  $G_{2,3} \{E_1, E_2\}$ , must be conjugate within the  $G_{M(s)}$  (and in fact, in this case  $D_{-1}$  belongs to the  $G_{M(s)}$ , in its real or second imaginary form, according as  $s = 12h + 1$  or  $12h + 11$ ).

(2°) We inquire under what groups  $G_{\infty t}$  ( $t \geq 1$ ) of the  $G_{2M(n)}$  and of the  $G_{M(n)}$  respectively this  $G_{\infty} \{E_1, E_2, E_3\}$  (12) is self-conjugate. Since under the  $G_{\infty}$  itself the various system  $\{E_1, E_2\}$  are conjugate, we seek a substitution  $U$  which transforms the system  $\{E_1, E_2\}$  and the group  $G_{\infty} \{E_1, E_2, E_3\}$  each into itself.  $U = I$  and  $U = D_{-1}$  are the only substitutions transforming  $\{E_1, E_2\}$  into itself.  $U = D_{-1}$  transforms  $G_{\infty} \{E_1, E_2, E_3\}$  into itself, if and only if  $q = 5$ . Now for  $q = 5$

$$U = D_{-1} = (-1, 0; 0, 1) = (-\sqrt{\rho}, 0; 0, \sqrt{\rho})$$

belongs to the  $G_{2M(n)}$  in its real or its second imaginary form, and to the real  $G_{M(n)}$ , but not to the second imaginary  $G_{M(n)}$ . Hence for  $q > 5$ , the  $G_{\infty} \{E_1, E_2, E_3\}$  is within the  $G_{2M(n)}$  or the  $G_{M(n)}$  self-conjugate exactly under itself, while, for  $q = 5$ , the  $G_{\infty} \{E_1, E_2, E_3\}$  is within the  $G_{2M(n)}$  self-conjugate under the  $G_{120} \{E_1, E_2, E_3, D_{-1}\}$  of the type  $G_{2M(n)}$  (§ 13, 5° B, 6° II;  $q = 5, k = 1$ ), and within the  $G_{M(n)=M(n^*)}$  under this  $G_{120}$  or under itself, according as  $n$  is even or odd, i. e., according as  $-3$  is or is not a square in the  $GF[5^n]$ .

(3°) Hence finally for all cases:

[Case  $q = 5$ .]<sup>22</sup> In the  $G_{M(n=5^n)}$  there are  $M(s)/60$  icosahedron groups  $G_{\infty} = G_{M(n)}$  conjugate within the  $G_{2M(n)}$  and forming within the  $G_{M(n)}$  two systems or one system of conjugate groups according as  $n$  is even or odd.

[Case  $q \neq 5$ .]<sup>24</sup> In the  $G_{M(n)} (s = q^n = 5k \pm 1)$  there are  $M(s)/(30; 60)$  icosahedron groups  $G_{\infty}$  conjugate within the  $G_{(3,1)M(n)}$  and forming within the  $G_{M(n)} (2; 1)$  systems of  $M(s)/60$  conjugate groups.

Of course, whenever within the  $G_{M(n)}$  ( $q > 2$ ) there are two systems, one obtains the representative of the second system by transforming by  $D_p$  the representative of the first system, viz., of the (one) system within the  $G_{2M(n)}$ .

§ 17. THE ABSTRACT GROUP  $G_{M(n)}$  IS HOLOEDRICALLY ISOMORPHIC TO THE (DOUBLY TRANSITIVE) SUBSTITUTION-GROUP  $G_{M(n)}^{s+1}$  OF DEGREE  $s+1$  AND TO A GROUP  $G_{M(n)}^D$  OF LOWER DEGREE  $D$  ( $D < s+1$ ) ONLY<sup>25</sup> IN THE CASES  $s = q^n = 5, 7, 11, 3^2$ ;  $M(s) = 60, 168, 660, 360$ ;  $D = 5, 7, 11, 6$  (WHEN THE  $G_{M(n)}^D$  IS RESPECTIVELY 3-, 2-, 2-, 4-PLY TRANSITIVE).

For  $n = 1$ , the theorem of the caption is a celebrated theorem due to Galois and Gierster.

(1°) For  $s = 2, 3$  the theorem is evident.

The  $G_{M(n)} (s \neq 2, 3)$  is simple (§ 8). A substitution-group  $G_{M(n)}^D$  is simple, only if it is (holoedrically isomorphic, or, say, for brevity) equivalent to its transitive constituents. We consider then, first, the *transitive*  $G_{M(n)}^D$  equivalent to our  $G_{M(n)}$ . This equivalence implies the existence in the  $G_{M(n)}$  of a (complete) system of  $D$  conjugate groups  $G_{\alpha} (D\Omega = M(s))$  (which have no common subgroup—and they do not, since

<sup>22</sup> Cf. § 13, 9°.

<sup>24</sup> Cf. § 13, 9° for  $q = 2, 3$ .

I communicated to Mr. Fricke in a letter dated February 17, 1894.

<sup>25</sup> This theorem (apart from the parenthetical remarks)

the  $G_{M(s)}$  is simple), and, conversely, if the  $G_{M(s)}$  contains such a system of  $D$  groups  $G_\alpha$ , then the operators of the  $G_{M(s)}$  *qua* transformers of these groups determine a  $G_{M(s)}^D$  equivalent to our  $G_{M(s)}$ .

We find that the  $s+1$  conjugate groups  $G_{M(s)}^{(s)} \frac{s(s-1)}{2+1} = \frac{M(s)}{(s+1)}$  (§ 12) of the  $G_{M(s)}$  ( $s \neq 2, 3$ ), which lead to the equivalent doubly transitive  $G_{M(s)}^{s+1}$ , have higher order than any other groups  $G_\alpha$  of the  $G_{M(s)}$ , for  $s > 11$ . Of the groups  $G_{M(s)}$  ( $s \leq 11, s \neq 2, 3$ ) subgroups of respectively (equal or, in fact), higher order occur in the cases  $s = 5, 7, 11, 3^2$ , viz., respectively, the one system of 5  $G_{12}$  (§ 15), the two systems of 7  $G_{24}$  (§ 15), the two systems of 11  $G_{60}$  (§ 16), the two systems of 6  $G_{60}$  (§ 16). Thus we have the  $G_{M(s)}^s$ , the  $G_{M(7)}^7$ , the  $G_{M(11)}^{11}$ , the  $G_{M(3^2)}^{3^2}$  of the caption theorem — for the two  $G_{M(7)}^7$  are similar, since under transformation by  $D_7$  the  $G_{M(s)}$  is invariant and the two systems of conjugate groups interchange, and similarly for the two  $G_{M(11)}^{11}$  and for the two  $G_{M(3^2)}^{3^2}$ .

The degree of an *intransitive*  $G_{M(s)}^D$  equivalent to the  $G_{M(s)}$  is the sum of the respective degrees of its two or more transitive constituents. It is clear that every such intransitive  $G_{M(s)}^D$  is of degree  $D > s+1$ .

(2°) The  $G_{M(3)=60}^5$  and the  $G_{M(3^2)=260}^6$  are of course respectively the triply, quadruply transitive alternating groups  $G_{120}^5, G_{160}^6$  (cf. also § 1, 5°).

The  $G_{M(7)=168}^7$  is doubly transitive. — For in the  $G_{M(7)}$  the 7 conjugate octahedron groups  $G_{24}'$  (of either system) are the elements of a triple system  $\Delta_7'$ . — For the  $G_{M(7)}$  has two systems of 7  $G_{24}$  (§ 15) —  $G_{24}', G_{24}''$  — and two systems of 7 four-groups  $G_{12}$  (§ 9) —  $G_{12}', G_{12}''$ . An octahedron  $G_{24}, e. g.$ , the  $G_{24,M(3)}$  (§ 1, 5°), contains a self-conjugate four-group  $G_{12}'$  and a system of 3 conjugate four-groups  $G_{12}''$  (§ 9). Every  $G_{24}'$  contains 3  $G_{12}''$  conjugate in itself and so in the  $G_{M(7)}$  — say 3  $G_{12,1}'$ . And every  $G_{12,1}'$  lies in 3  $G_{24}'$ . The 7  $G_{12,1}'$  determine then 7 triples of  $G_{24}'$ , in fact<sup>20</sup> a triple system  $\Delta_7'$ . — Now all triple systems  $\Delta_7$  are similar. A  $\Delta_7$  has a (largest) group  $G_{168}'$  ( $\Delta_7'$ ) of order 168 and doubly transitive. Our  $G_{M(7)}^7$  on the 7  $G_{24}'$  is then the group  $G_{168}'$  ( $\Delta_7'$ ) leaving invariant the triple system  $\Delta_7'$  in the 7  $G_{24}'$ .

The  $G_{M(11)=660}^{11}$  is doubly transitive. — For the  $G_{M(11)}$  contains two systems of 55 dihedron  $G_{12}$  and two systems of 11 icosahedron  $G_{60}$  (§§ 9, 16); every  $G_{60}$  contains 10 conjugate  $G_{12}$ ; every  $G_{12}$  lies in two conjugate  $G_{60}$  (§ 16, 1°, footnote). Hence the  $G_{M(11)}$  is transitive on the 11  $G_{60}'$ ; a  $G_{60}'$  is transitive on its 10  $G_{12}'$  and so on the remaining 10  $G_{60}'$ . Thus the  $G_{M(11)}$  is doubly transitive on the 11  $G_{60}'$ . (Its order precludes its being triply transitive.)

<sup>20</sup> In the  $G_{24,M(3)}$  the three  $G_{12,1}''$  contain each a (different) substitution  $V_3'$  of the  $G_{12,1}'$  (§ 9). A  $G_{12,1}'$  and the  $G_{12,1}''$  lie in and determine a dihedron  $G_{24}$ , containing two substitutions  $V_2'$ , three  $V_3'$ , and two substitutions  $V_4'$  of period 4, besides the identity. Two  $G_{12,1}'$  generate a group containing such a  $G_{24}$  and another substitution  $V_3'$  — and so the whole

$G_{24,M(3)} = 24$ . — Hence in the  $G_{M(7)}$  two  $G_{24}'$  having a  $G_{12,1}'$  in common do not have two  $G_{12,1}''$  in common. Thus no two triples of  $G_{24}'$  have two  $G_{12,1}''$  in common. The 7 triples contain then the 7.6 pairs of  $G_{24}'$  without duplication and without omission, and hence constitute a triple system  $\Delta_7'$ .

**A CONTRIBUTION TO THE THEORY OF  
GLACIAL MOTION**



# A CONTRIBUTION TO THE THEORY OF GLACIAL MOTION

T. C. CHAMBERLIN

THE key to this study is the thesis that a glacier is a mass of crystalline rock—the purest and simplest type of crystalline rock known—since it is made up of a single mineral, of simple composition and rare purity, which never appears in a solid state except in the crystalline form.

## THE GROWTH AND CONSTITUTION OF A GLACIER

The origin and history of a glacier is little more than the origin and aggregate history of the crystals that compose it. The fundamental conception of a glacier is therefore best obtained by tracing the growth of its constituent crystals. A basal fact ever to be kept in mind is that water in the solid form is always controlled by crystalline forces. When it solidifies from the vapor of the atmosphere, it takes the form of separate crystals (Plate I, Fig. 1). Perfect forms are developed only when the flakes fall quietly through a saturated atmosphere which allows them to grow as they descend. Under other conditions, the crystals are imperfect in growth and are mutilated by impact. But, however modified, they are always crystals. The molecules are arranged on the hexagonal plan, and the assemblage is controlled by a strong force, as the expansive power of freezing water shows. Once the definite crystalline arrangement is established, the molecules can be displaced only by the expenditure of a very notable amount of energy.

Snow crystals often continue to grow so long as they are in the atmosphere; but if they pass through an undersaturated stratum of air, or a stratum whose temperature is above 32° F., they suffer from evaporation or melting. When they reach the ground, the processes of growth and decadence continue, and the crystals grow or diminish according to circumstances.

*A glacier is a colossal aggregation of crystals grown from snowflakes to granules of much greater sizes.*

The microscopic study of new-fallen snow reveals the mode of change from flakes to granules. The slender points and angles of the former yield to melting and evaporation more than the more massive central portions, and this change probably illustrates a law of vital importance. It may often be seen that the water melted from the periphery of a flake gathers about its center, and if the temperature be right, it freezes there. This is a first step toward the pronounced granulation of snow, illustrations of which are familiar in snow that has lain long on the ground. If measured systematically from day to day, the larger granules taken from beneath the surface of coarse-grained snow are found to be growing. In a series of experiments<sup>1</sup> to determine the

<sup>1</sup> Carried out by Mr. C. E. Peet and Mr. E. C. Perisho, under the direction of the author.



law of growth, it was found that when the temperature of the atmosphere was above the melting-point, the growth was appreciably more rapid than when the air was colder, but that there was, on the average, *an increase under all conditions of temperature*. A portion of this average increase of the larger granules appears to come from the diminution and destruction of the smaller ones, for the total number of granules steadily diminishes. A portion of the growth doubtless comes from the moisture of the atmosphere which penetrates the snow, and another portion from the moisture derived from surface melting; but beneath the surface of a large body of snow the growth of the large granules probably takes place chiefly at the expense of the small ones. To follow the process, it should be noted that the free surface of every granule is constantly throwing off particles of vapor; that the rate at which the particles are thrown off is dependent, among other things, on the curvature of the surface being greater the sharper the curve; that the surfaces of the granules are at the same time liable to receive and retain molecules thrown from other granules; and that, other things being equal, the retention of particles also depends on the curvature of the surface, but in the reversed sense, the less curved surface retaining more than the sharply curved one. Under these laws, it is obvious that the larger granules of smaller surface curvature will lose less and gain more, on the average, than the smaller granules of greater curvature. It follows that the larger granules will grow at the expense of the smaller. It is also to be noted that, other things being equal, small granules melt more readily than large ones, and that where the temperature is nicely adjusted between melting and freezing the smaller may lose while the larger gain.

Another factor that enters into the process is that of pressure and tension. The granules are compressed at the points of contact, and put under tension at points not in contact; and the pressure and tension are, on the average, likely to be relatively greatest for the smallest granules. Tension increases the tendency to evaporation, and adds its effects to surface curvature. The capillary spaces adjoining the points of contact probably favor condensation. Ice expands in crystallizing, and pressure reduces the melting-point, while tension raises it. The effect of this is slight ( $0.0075^{\circ}$  C. per atmosphere), and it probably plays little part in glacial action; but it is to be correlated with the much more important fact that *compression produces heat*, which may raise the temperature of the ice to the melting-point, while tension reduces the temperature to or below freezing. There is therefore a tendency for the ice to melt at the points of contact and compression, and for the water so produced to refreeze at adjacent points where the surface is under tension. This process becomes effective beneath a considerable body of snow, and here the granules gradually lose the spheroidal form assumed in the early stages of granulation, and become irregular polyhedrons interlocked into a more or less solid mass.

A third factor is also to be recognized, though its effectiveness is unknown. Under severe wind pressure, air penetrates porous bodies with appreciable facility. The "breathing" of soils and the curious phenomena of "blowing wells" and "blowing

caves" illustrate the effective penetration and extrusion of the air under variations of barometric pressure. In the snowfields, and in the more granular portions of glaciers near their heads, the porosity is doubtless sufficient to permit the appreciable penetration of the atmosphere. During a part of the time the probable effect is condensation of moisture from the air within the ice, and during another part, evaporation from the exposed surfaces of the granules within the ice. These alternating processes are attended by oscillations of temperature. While the balance between loss and gain of substance to the mass may be immaterial, the oscillating nature of the process and the fluctuations of temperature are probably favorable to the transference of molecules from point to point, and hence favorable to granular change.

Whether these processes furnish the total, or even the essential, explanation of the process or not, the observed fact is that there are all gradations from snowflakes and pellets into granular névé, and thence into glacier granules (*Gletscherkörner*), the last reaching the size of filberts and walnuts and beyond. The state of aggregation varies from the early, slightly coherent, granular stage, where the grains are small and spheroidal, and loosely attached, to the ice stage, where the cohesion has become strong through the interlocking growths of the large granules. Even when the mass has become seemingly solid ice, sufficient space is usually left between the granules to give a dispersive reflection to light which imparts to glacier ice its distinctive whitish color.

#### THE ARRANGEMENT OF THE CRYSTALLINE AXES

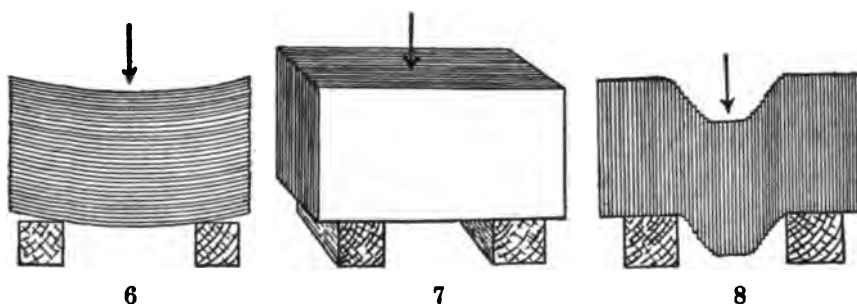
The most radical difference between glacier ice and ice formed directly from water is in the orientation of the crystals. In ice formed on undisturbed water the bases of the crystals are at the surface, and their principal axes are vertical, as shown by Mügge.<sup>1</sup> As they grow, the crystal prisms extend downward. This gives a columnar or prismatic structure to the ice, well shown when it is "honeycombed" by partial melting. In a glacier, on the other hand, the axes of the crystals, starting from snowflakes, lie originally in every direction, according to the accidents of their fall; and as the snow develops into ice, the principal axes of the crystals continue, in the main, to lie in every direction. Hence glacier ice, unlike pond ice, cannot usually be split along definite planes, except where cleavage planes are subsequently developed by extraneous agencies.

While the crystals of a glacier usually have their principal axes in various directions, there appears to be a tendency for them to approach parallelism in certain positions, especially in the lower part of a glacier near its terminus. Observations on this point are not so full and critical as could be desired, but it is probable that the parallel orientation is partly general and due to the vertical pressure of the ice, and partly special and local, and connected with the shearing planes and foliation (illustrated in Plate I, Figs. 2-4; Plate II, Figs. 5, 9-11; Plate III, Figs. 12-14).

The bearing of this partial parallelism of the crystals on shearing and foliation is

<sup>1</sup> Mügge, "Ueber die Plasticität der Eiskrystalle," *Neues Jahrbuch für Mineralogie*, etc., 1895, Bd. II, p. 211.

supposed to reside in the fact that a crystal of ice is made up of a series of plates arranged at right angles to the principal axis of the crystal. These plates may be likened to a pile of cards, the principal axis being represented by a line vertical to them. If a cube be cut from a large crystal of ice, it will behave much like a cube cut from the pile of cards. If the cube be so placed that its plates are horizontal, and if it be rested on supports at two edges and heavily weighted in the middle (Fig. 6), it will sag, the plates sliding slightly over one another so as to give oblique ends, but in this case the cube offers considerable resistance to deformation. If the cube be so placed that the plates stand on edge and stretch from support to support (Fig. 7), it will offer very great resistance to deformation; but if the plates be vertical and transverse to the line adjoining the supports (Fig. 8), the middle portion will sag under moderate weighting by the sliding of the plates on one another, and in a comparatively



FIGS. 6, 7, 8.—Diagrammatic figures to illustrate the method of deformation of crystals. (Adapted from Mägge.)

short time the middle portion may be pushed entirely out, dividing the cube. These properties have been demonstrated by McConnell<sup>3</sup> and Mägge,<sup>4</sup> and they appear to throw light on certain phases of the action of glaciers that are most pronounced in their basal parts and are best illustrated in arctic glaciers.

#### GLACIER TEMPERATURES

The temperature of glacier ice may range *downward* from the freezing-point of water, much as that of other solid portions of the earth's surface, but it has a fixed *upper* limit at 0° C., because all the heat it receives tending to raise its temperature above that point is converted into the latent form by the melting of the ice. The range of temperature is greatest at the surface, where it varies from 0° C. in the summer to the coldest temperature of the region where the ice occurs. Beneath the surface the range of temperature is more restricted, and increasingly so with increasing depth.

The variation of temperature at the surface is due primarily to the varying temperature of the air. During the cold season a wave of low temperature (the *winter* wave), starting at the surface, penetrates the ice, and during the warm season a wave of

<sup>3</sup> McCONNELL, "On the Plasticity of Glaciers and Other Ices," *Proceedings of the Royal Society*, Vol. XLIV (1888), pp. 331-67 (with D. A. Kidd); Vol. XLVIII (1890), pp. 249-60; Vol. XLIX (1891), pp. 323-43.

<sup>4</sup> MÄGGE, *loc. cit.*

higher temperature (the *summer* wave) takes the same course. The day and night waves and other minor variables are, for present purposes, negligible.

*The winter wave.*—There are but few observations on the internal temperatures of glaciers during the winter season, but it seems certain that the winter wave diminishes rapidly downward and dies out below, much as does the winter wave which affects land surfaces not covered with ice. Conduction alone considered, the temperature of the ice where the winter wave dies out should correspond, approximately, to the mean annual temperature of the region, provided that temperature is below the melting-point of ice.

Assuming that in the high altitudes and high latitudes where glaciers abound the temperature of the surface may average about  $-25^{\circ}$  C. for the winter half of the year—which is about the case for north Greenland, Spitzbergen, and Franz Josef Land—and that the conductivity of the ice in the C. G. S. system is 0.005, the temperature would be lowered appreciably only about forty feet below the surface at the close of the period, conduction only being considered.<sup>5</sup> How far the internal temperature may be influenced by air forced through it by winds and by variations of the barometer is not known, and cannot well be estimated. The wave of low temperature descending from the surface in winter would probably become inappreciable before reaching a depth of sixty feet. At this depth the temperature should be about  $-10^{\circ}$  C., or near the mean annual temperature of the region.

*The summer wave.*—The summer wave in ice follows the analogy of the summer wave of ordinary earth much less closely, because of the melting of the ice. On this account, the summer wave is bifold. The one part travels downward by conduction; the other, by the descent of water. The one has to do primarily with the temperature before the melting-point of ice is reached; the other, with the temperature after that point is reached. The first conforms measurably to the warm wave affecting other solid earth-matter, while the second is governed by laws of its own. After the surface portion of the ice is brought to the melting temperature, the additional heat which it receives melts the ice and is transformed from *sensible* into *potential* heat. Ice charged with water is *potentially* but not *sensibly* warmer than ice which has just reached the melting temperature.

The warm wave of conduction dies out below, like the cold wave. The warm wave descending by the flow of water stops where the freezing temperature of water is reached. In regions where the average temperature is below freezing, the water wave does not descend so far as the wave of conduction, since the latter descends below the zone where the melting temperature is found.

The foregoing considerations warrant the generalization that glaciers normally consist of two zones: (1) an outer or upper zone of fluctuating temperature, and (2) an under zone of nearly constant temperature. The under zone does not exist where the thickness of the ice is less than the thickness of the zone of fluctuating temperature.

<sup>5</sup> Computation by E. D. K. Leffingwell.

This may be the case in very thin glaciers in very cold regions, and it may be true of the thin ends and edges of all glaciers.

*The temperature of the bottom.*—The internal heat of the earth is slowly conducted to the base of a glacier, where it melts the ice at the estimated average rate of about one-fourth of an inch per year. The temperature of melting is a little below  $0^{\circ}$  C., since pressure lowers the melting-point at the rate of  $0.0075^{\circ}$  C. for one atmosphere of pressure. At the bottom of a mile of ice, therefore, the melting temperature is about  $-1^{\circ}$  C. It is probable that in all thick glaciers the temperature of the bottom is constantly maintained at the melting-point. This may be indicated by the streams which issue from beneath glaciers during the winter, though the indication is hardly decisive, since the issuing waters may be derived partly or wholly from the rock beneath. In glaciers or parts of glaciers so thin as to lie wholly within the zone of fluctuating temperature, the temperature of the bottom is obviously not constant.

*The internal temperature of the ice.*—The range of temperature of the surface of a glacier has already been shown to lie between a maximum of  $0^{\circ}$  C. and the minimum temperature of the region where the glacier occurs. Lower, in the zone of fluctuating temperature, the range is less, and where the zone of fluctuating temperature passes into the zone of constant temperature, variation ceases. The thickness of the zone of fluctuating temperature varies with the temperature of the region where the glacier occurs, being greatest where the winters are coldest.

The range of temperature within the zone of constant temperature, in the case of all glaciers except thin ones in very cold regions, is from the mean annual temperature of the region at the top of the zone (provided this is not above the melting-point of ice) to the melting temperature of the ice at the bottom. Within these limits, the differences of temperature may be great or slight.

If we consider only the effects of the external seasonal temperatures and the internal heat of the earth, it appears that all the ice in the zone of constant temperature in the lower end of a typical alpine glacier should have a constant melting temperature; for the average temperature of regions where the ends of such glaciers occur is usually above  $0^{\circ}$  C., and this determines a temperature of  $0^{\circ}$  C. (or a little less) at the top of the zone, while a melting temperature is maintained at the bottom by the earth's interior heat. In thin glaciers of very cold regions, where the zone of constant temperature has relatively slight thickness, or lies wholly below the ice, the low temperature descending from the surface may so far overcome the effect of internal heat as to keep the bottom of the ice at a freezing temperature. In all other cases the ice at the bottom of the under zone has a melting temperature, while that above is probably colder.

In the higher altitudes and in the polar latitudes, where glaciers are chiefly generated, the mean annual temperature of the surface is *usually below the melting-point of ice*. Here the temperature of the ice between the top and the bottom of the zone of constant temperature must, on the average, be below the melting-point, unless

enough heat is generated in the interior of the ice to offset the effect of the temperature above. For example, where the mean annual temperature is  $-6^{\circ}$  C., as in certain high altitudes and high latitudes, the mean temperature in the zone of constant temperature should range from  $-6^{\circ}$  C. above to  $0^{\circ}$  C. (or a little less, because of pressure) below; *i. e.*, it should average about  $-3^{\circ}$  C. Under these conditions, all the ice in the zone of constant temperature, except that at its bottom, is permanently below the melting-point; but it is perhaps worthy of special note that much of it is but little below. In alpine glaciers the part of the ice affected by this constant freezing temperature is presumed to be chiefly that which lies beneath the snowfields. In polar glaciers the low temperature probably prevails beneath the surface not only throughout the great ice-caps, but also in the marginal glaciers which descend from them.

From these theoretical considerations we may deduce the generalization that in the zone of constant temperature *within the area of glacial growth* the temperature of the ice is generally *below* the melting-point, while *within the area of wastage* the temperature of the corresponding zone is generally *at* the melting-point.

*Compression and friction as causes of heat.*—The foregoing conclusions are somewhat modified by certain other sources of heat. The compression arising from gravity, and the friction developed where there is motion, are causes of heat. Since friction occurs only when motion takes place, the heat which it generates is secondary and may, for present purposes, be neglected. Compression not only lowers the melting-point slightly, but it *produces heat at the point of compression*. Where the ice is granular, the compression, due to weight, takes place at the contacts of the grains. At intermediate points the pressure tends to cause them to bulge, and this has the effect of lowering the temperature of the bulging areas. If therefore the compression be considerable, the granules may be warmed to the melting-point where they press each other, while at other points their temperature may be lower. In this case melting will take place at the points of compression, and the moisture so produced will be transferred to the adjacent parts of the granule and immediately refrozen. Melting at the points of compression would result in some yielding of the mass, and in some shifting of the pressure to new points where compression and melting would again take place. Thus the melting, the refreezing, and the attendant movement might go on until the limits of the power of gravity in this direction were reached. From considerations already adduced it would appear that the temperature in some parts of every considerable body of ice must be such as to permit these changes. This dynamic source of heat may modify the theoretical deductions drawn above from atmospheric and internal influences.

*Summary.*—If the foregoing generalizations be correct, (1) the surface of a glacier is likely to be melted during the summer; (2) its immediate bottom is slowly melting all the time, unless the thickness of the ice be less than the thickness of the zone of annual variation or of permanent freezing temperature; (3) its subsurface portion

*in the zone of waste* is generally melting, owing to descending water, compression, and friction; while (4) its subsurface portion *in the zone of growth* is probably below the melting-point, except as locally brought to that temperature by compression and friction, and at the bottom by conduction from the rock beneath.

#### FUNDAMENTAL AS DISTINGUISHED FROM AUXILIARY CAUSES OF MOTION

Since there must be motion in the area of growth to supply the loss in the area of waste, the *fundamental* cause of motion must probably be operative in bodies of ice the mean temperature of which is below the melting-point, *unless the internal sources of heat are considerable*. This fundamental cause does not exclude the *co-operation* of causes that work only (1) at the melting temperature, (2) where the ice is bathed with water, or (3) in the plane of contact between wet ice above and dry ice below. These may be *auxiliary* causes which abet the fundamental one in producing the more rapid movement of the warm season, or in bringing about especially rapid motion in situations where there is abundant water, or in inducing the shearing which is such a remarkable feature of arctic glaciers.

#### THE PROBABLE FUNDAMENTAL ELEMENT IN GLACIAL MOTION

*Melting and refreezing.*—It has just been pointed out that the *initial* or *fundamental* cause of glacial motion must be operative at the heads of glaciers where the temperature is lowest and the material most loosely granular. In this condition, there is reason to believe that motion takes place between the grains, rather than by their distortion through the displacement of their laminae. The fact that the granular structure is not destroyed, as it would be by the indefinite sliding of the crystal plates over one another, sustains this view. The inference is that the gliding planes play a notable rôle in glacial movement only in the basal parts of the lower ends of glaciers, where the greatest thrusts are developed, and where the granules have become largest and most completely interlocked. At the heads of glaciers, where motion is initiated, there may be great downward pressure, but not vigorous thrusts from behind, and probably only moderate thrusts developed within the body itself. *There seems therefore no escape from the conclusion that the primal cause of glacial motion is one which may operate even under the relatively low temperature, the relatively dry condition, and the relatively granular texture which affect the heads of glaciers.*

These considerations lead to the view that movement takes place by the minute individual movements of the grains upon one another. While they are in the spheroidal form, as in the *névé*, this would not seem to be at all difficult. They may rotate and slide over one another as the weight of the snow increases; but as they become interlocked by growth, both rotation and sliding must apparently encounter more resistance. The amount of rotary motion required of an individual granule is, however, surprisingly small, and the meltings and refreezings incident to shifting pressures and tensions, and to the growth of the granules, seem adequate to meet the require-

ments. To account for a movement of three feet per day near the end of a glacier six miles long, the mean motion of the average granule relative to its neighbor is, roundly, one ten-thousandth of its own diameter per day, or one diameter in ten thousand days; in other words, it *changes its relations to its neighbors* to the extent of its diameter in about thirty years. A change of so great slowness under the conditions of granular alteration can scarcely be thought incredible, or even improbable, in spite of the interlocking which the granules may develop. The movement is supposed to be permitted chiefly by the temporary passage of minute portions of the granules into the fluid form at the points of greatest compression, the transfer of the moisture to adjoining points, and its resolidification. The points of greatest compression are obviously just those whose yielding most promotes motion, and a successive yielding of the points that come in succession to oppose motion most (and thus to receive the greatest stresses) permits continuous motion. It is merely necessary to assume that the gravity of the accumulated mass is sufficient to produce the minute temporary liquefaction at the points of greatest stress, the result being accomplished, not so much by the lowering of the melting-point, as by the development of heat by pressure.

This conception of glacial "flowage" involves only the *momentary liquefaction of minute portions of the mass*, while the ice, as a whole, remains rigid, as its crystalline nature requires. Instead of assigning a slow viscous fluidity, like that of asphalt, to the *whole* mass, which seems inconsistent with its crystalline character, it assigns a free fluidity to a succession of particles that form only a minute fraction of the whole at any instant.

This conception is consistent with the retention of the granular condition of the ice, with the heterogeneous orientation of the crystals (in the main), with the rigidity and brittleness of the ice, and with its strictly crystalline character—a character which a viscous liquid does not possess, however much its high viscosity may make it resemble a rigid body.

*Accumulated motion near the end of a glacier.*—However slight the relative motion of one granule on its neighbor, the granules in any part of a glacier partake in the accumulated motion of the parts nearer the source, and hence all are thrust forward. Herein appears to lie the distinctive nature of glacial movement. Each part of a stream of water feels the hydrostatic pressure of neighboring parts (theoretically equal in *all* directions) and the momentum of motion, but not the rigid thrust of the mass behind. Lava streams are good types of viscous fluids flowing in masses comparable to those of glaciers, on similar slopes, and, in their last stages, at similar rates; but their special modes of flow and their effects on the sides and bottoms of their paths are radically different from those of glaciers. Forceful abrasion, and particularly the rigid holding of imbedded stones while they score and groove the rock beneath, are unknown in lava streams, and are scarcely conceivable. There is, so far as I know, neither experimental nor natural evidence that any typical viscous body, in flowing over a rugose bottom, picks up fragments and holds them as graving tools in its base



so fixedly as to cut deep long straight grooves in the hard bottom over which it flows. It would seem that competency to do this peculiar class of work, which is distinctive of glaciers, should be demonstrated before the viscous theory of glacial movement is accepted as even a provisional working hypothesis. Quite in contrast with viscous movement, it is conceived that a glacier is thrust forward rigidly by internal elongation, shears forcibly over its sides and bottom, and leaves its distinctive marks upon them.

#### AUXILIARY ELEMENTS IN GLACIER MOTION

*Shearing.*—In the lower portion of a glacier, where normally the thrusts are greatest, the granules fewest, and their interlocking most intimate, shearing appears to take place within the ice itself. This is illustrated by Figs. 2-5, and 9-14 (Plates I, II, and III). The shearing results in the foliation of the ice and in the forcing of débris between the sheared layers. Thus the ice becomes loaded in a special baso-englacial fashion (see particularly Fig. 14).

Within the zone of shearing, it is probable that the gliding planes of the crystals come into effective function. It is thought that the combined effect of the vertical pressure, the forward thrust, and the basal drag of the ice may be to increase the number of granules whose gliding planes are parallel to the glacier's bottom. At any rate, Drygalski reports<sup>6</sup> that there is a tendency to such an arrangement in the basal portion of the Greenland glaciers at their borders. It is conceived that where strong thrusts are brought to bear upon such a mass of granules, those whose gliding planes are parallel to the direction of thrust are strained with sufficient intensity to cause the plates to slide over one another, while those which are not parallel to the direction of thrust are either rotated into parallelism—when they also yield—or are pressed aside out of the plane of shear.<sup>7</sup> Shearing is observed to occur chiefly where the ice below the plane of shearing is protected more or less from the force of the thrust. It perhaps also occurs where the basal ice becomes so overloaded with débris that it is incapable of ready movement.

It is also probable that sharp differential strain and shearing are developed at the level where the surface water of the warm season descending into the ice reaches the zone of freezing. The expansion of the freezing water at the upper limit of the frozen zone may cause the layer affected by it to shear over that below. As the level of freezing is lowered with the advance of the warm season, the zone of shearing also sinks. This may be regarded as an auxiliary agency of shearing of application to the upper horizons.

<sup>6</sup> *Grönland-Expedition der Gesellschaft für Erdkunde zu Berlin*, 1891-93, Bd. I, pp. 491 ff.

<sup>7</sup> A series of experiments on compacted snow was conducted during a winter season to test this. About three-fourths of the granules in the tract of greatest strain took on a common orientation. The rigid resistance of the snow-ice proved very great, and differential pressures, increased to the breaking strength of the apparatus, nearly

one thousand pounds per square inch, were insufficient to produce, in the time given, a definite shear-plane of the Greenland type, but the approximation attained lends strength to the belief that a new series of experiments, contemplated but as yet unexecuted, in which the rigidity of the snow-ice shall be fully anticipated and the time requisite for molecular change provided for, will fully sustain the view here entertained.

*High temperature and water.*—In the zone of waste, a higher temperature and more water lend their aid to the fundamental agencies of movement; and there is need for these aids to promote proportionate movement, for here the granules are more intimately interlocked, and the ice more compact, and inherently more solid and rigid. The average temperature here is, however, near the melting-point, and during the warm season the ice is bathed in water, so that the necessary changes in the crystals are facilitated, and movement apparently takes place even more readily than in the less compact granular portion of lower temperature and drier state. The extraordinary rate of movement of certain tongues of ice in some of the great fiords of Greenland is probably due to the convergence of very thick, slow-moving ice from the interior, into basins leading down to the fiords. Into the same basins a large amount of surface water is concentrated at the same time, with the result that the thick ice, bathed with water, and having a high gradient, develops unusual velocity during the warm season.

*Application.*—By a studious consideration of the co-operation of the auxiliary agencies with the fundamental ones, the peculiarities of glacial movement may apparently be explained. In regions of intense cold where a dry state and low temperature prevail, as in the heart of Greenland, the snow-ice mass may accumulate to extraordinary thickness, for the work of movement seems to be thrown almost wholly upon compression, with the slight aid of molecular changes due to internal evaporation and allied inefficient processes. Since the temperature in the upper part of the ice is very adverse, the compression must be great before it becomes effective in melting the contacts of the granules of ice, and hence the great thickness of the mass antecedent to much motion. Similar conditions more or less affect the heads of alpine glaciers, though here the high gradients favor motion with less thickness of ice; but in the lower reaches of alpine glaciers, where the temperatures are at or near the melting-point, and the ice is bathed in water, movement may take place in ice which is thin and relatively compact.

If the views here presented are correct, there is also near the end or edge of a glacier the co-operation of rigid thrust from behind with the tendency of the mass to move on its own account. The latter is controlled by gravity, and conforms in its results to the laws of liquid flow. The former is a derived factor, and is a mechanical thrust. This thrust is different from the pressure of the upper part of a liquid stream on the lower part, because it is transmitted through a body whose rigidity is effective, while the latter is transmitted on the hydrostatic principle of equal pressure in all directions.

#### CORROBORATIVE PHENOMENA

The conception of the glacier and its movement here presented explains some of the anomalies that otherwise seem paradoxical. While a glacier, in a sense, "flows" over a surface, it often cuts long deep furrows in firm rock (Plate III, Fig. 15). It is difficult to explain this if the ice be so yielding as to flow under its own weight on a surface which is almost flat. If the mass is really viscous, its hold on its imbedded

débris should also be viscous, and a boulder in the bottom should be rotated in the yielding mass when its lower point catches on the rock beneath, instead of being firmly held while a deep groove is cut. This is the more to the point since viscous fluids flow by a partially rotatory movement. If, on the other hand, the ice is always a rigid body, which yields only as its interlocking granules change their form by loss and gain, a rigid hold on the imbedded rock at some times, and a yielding hold at others, is intelligible, for, on this view, the nature of its hold is dependent on the temperature and dryness of the ice. Stones in the base of a glacier may be held with very great rigidity when the ice is dry and cold, scoring the bottom with much force; while they may be rotated with relative ease when the ice is wet and warm. In short, the relation of the ice to the boulders in its bottom varies radically according to its dryness and temperature. *A dry glacier is a rigid glacier. A dry glacier is necessarily cold, and a cold glacier is necessarily dry.*

On the view here presented, a glacier should be more rigid in winter than in summer, and the whole thickness of the glacier should experience this rigidity chiefly at the ends and edges, where the relative thinness of the ice permits the low temperature to reach its bottom. The motion in these parts during the winter is therefore very small.

In this view may also be found an explanation of the movement of glaciers for considerable distances on upward slopes, even when the *surface*, as well as the base, is inclined backward. So far does this go that superglacial streams sometimes run for some distance *backward*—i. e., toward the heads of the glaciers—while in other places surface waters are collected into ponds and lakelets. Such a forward ascent of the surface of ice is not difficult to understand, if the movement be due to thrust from behind, or if it be occasioned by internal crystalline changes acting upon a rigid body; but it must be regarded as very remarkable, if the movement be that of a fluid body, no matter how viscous, for the length of the activity is sometimes several times the thickness of the ice. Crevassing, and other evidences of brittleness and rigidity, find a ready elucidation under the view that the ice is a really solid body at all times, and that its apparent fluency is due to the momentary fluidity of small portions of the mass assumed in succession as compression demands.

In addition to the considerations already adduced, it may be urged that a glacier does not flow as a stiff liquid, because its granules are not habitually drawn out into elongated forms, as are cavities in lavas and plastic lumps in viscous bodies. Flowage lines, comparable to those in lavas, are unknown in glaciers.

All this is strictly consistent with our primary thesis that a glacier is a crystalline rock of the purest and simplest type, and that it never has other than the crystalline state. This strictly crystalline character is incompatible with viscous fluidity.

## BIBLIOGRAPHY

- AGASSIZ, L. *Études sur les glaciers*, Neuchâtel, 1840.
- CHAMBERLIN, T. C. "Presidential Address before the Geological Society of America." *Bull. Geol. Soc. Am.*, Vol. VI (February, 1895), pp. 199-220.
- CHARPENTIER, J. DE. *Essai sur les glaciers et le terrain erratique du Bassin du Rhône*, Lausanne, 1841.
- CROLL, JAMES. "On the Physical Cause of the Motion of Glaciers." *Phil. Mag.*, Vol. XXXVIII (1869), pp. 201-6.
- DEELEY, R. M. "A Theory of Glacial Motion." *Phil. Mag.*, Vol. XXV (1888), pp. 136-64.
- DEELEY, R. M., AND FLETCHER, GEORGE. "The Structure of Glacier Ice and Its Bearings on Glacier Motion." *Geol. Mag.*, London, Decade 4, Vol. II (1895), pp. 152-62.
- DRYGALSKI, ERICH VON. *Grönland-Expedition d. Gesells. f. Erdkunde zu Berlin 1891-93*, Vol. I (1897).
- FORBES, J. D. *Occasional Papers on the Theory of Glaciers*, Edinburgh, 1859.
- GRAD, CH. "La constitution et le mouvement des glaciers." *Revue sci.*, 1872.
- HEIM, A. "On Glaciers." *Phil. Mag.*, Vol. XLI (1871), pp. 485-508; *Handbuch der Gletscherkunde*, 1885.
- HOPKINS, W. "On the Theory of the Motion of Glaciers." *Phil. Trans.*, 1862, p. 677; *Phil. Mag.*, Vol. XXV (1863), p. 224.
- HUGI, F. J. *Ueber das Wesen der Gletscher und Winterreise in dem Eismeer*, Stuttgart, 1842.
- MALLET, R. "The Mechanism of Glaciers." *Jour. Geol. Soc.*, Dublin, Vol. I, p. 317.
- "On the Plasticity of Glacier Ice." *Ibid.*, Vol. III (1845), p. 122.
- "On the Brittleness and Non-Plasticity of Glacier Ice." *Phil. Mag.*, Vol. XXVI, p. 586.
- McCONNEL, J. C. "On the Plasticity of an Ice Crystal." *Proc. Roy. Soc.*, London, Vol. XLVIII (1890), pp. 256-60; Vol. XLIX (1891), pp. 323-43.
- MOSELEY, H. "On the Cause of the Descent of Glaciers." *Brit. Ass. Rept.*, Part II (1860), p. 48; *Phil. Mag.*, Vol. XXXVII (1869), pp. 229, 363; Vol. XXXIX, p. 241; Vol. XLII, p. 138; Vol. XLIII, p. 38.
- MÜGGE, O. "Ueber die Plasticität der Eiskrystalle." *Nachr. K. Ges. d. Wiss.*, Göttingen, 1895, pp. 1-4.
- RENDU. "Théorie des glaciers de la Savoie." *Soc. Roy. Acad., Savoie, Mém. 1840* (in English, ed. by Geo. Forbes, London, 1874).
- RINK, H. J. *Danish Greenland*, 1877.
- THOMPSON, JAMES. "On the Plasticity of Ice as Manifested in Glaciers." *Roy. Soc. Proc.*, Vol. VIII (1857), pp. 455-58.
- TYNDALL, J. *Forms of Water*, New York, 1872.
- The Glaciers of the Alps*, London, 1861.
- TYNDALL, J., AND HUXLEY, T. H. "On the Structure and Motion of Glaciers." *Phil. Trans.*, Vol. CXLVII (1857), p. 327.

## EXPLANATION OF PLATES

## PLATE I

FIG. 1.—Snow crystals. (Photographed by W. A. Bentley.)

FIG. 2.—Terminal portion of Bryant Glacier, Inglefield Gulf, north Greenland, seen from the east, showing clean white ice above and stratified, discolored ice below, with a talus slope at the base. Height of vertical face above the gravel plain in front, about 140 feet. In this and

the following views the amount of *débris* in the discolored ice is much less than it appears to be, as it is greatly exaggerated to appearances by surface spreading.

FIG. 3.—Side view of middle portion of Bryant Glacier, showing the relation of the foliated *débris*-laden base of the ice to the clean white ice above.

FIG. 4.—Terminal portion of Bryant Glacier, seen from the west, showing the foliated structure of the basal ice.

#### PLATE II

FIG. 5.—A near view of the basal *débris*-bearing layers of the Bryant Glacier at a point where they are turned up into a nearly vertical attitude. The surface is much covered with *débris* derived from the ice above. This has been cut away from a portion to show the real amount of *débris* included in the ice, and its definite arrangement in layers.

FIGS. 6-8, text figures.

FIG. 9.—Portion of the end of South Point Glacier, Bowdoin Bay, Inglefield Gulf region, north Greenland, showing nearly pure ice above, but well foliated or stratified, and a boulder-set layer below, from whose melting the talus slope of *débris* is formed. The ice of the glacier underlies this talus. The talus slope appears to be nearly or quite stationary, and the chief horizon of differential ice motion is probably at the boulder-set layer at the top of the talus slope. The plain in front is formed of glacial wash.

FIG. 10.—Another portion of the end of South Point Glacier, showing the foliation of the ice brought out by shadows rather than by *débris*. The dark tract forming the lower part is ice covered with dirt from embraced *débris*. The base is a talus slope derived from this dark layer, but it is underlain by ice.

FIG. 11.—View of the edge of Krakokta Glacier—a tongue from the small ice-cap of Red Cliff Peninsula, Inglefield Gulf region, north Greenland, showing the strong individuality of the ice layers in the basal portion. The influence of *débris* in causing the layers to stand forth is here not very obvious, and the question of the individuality of motion of the layers is more open.

#### PLATE III

FIG. 12.—View of the terminal face of the Tucktoo Glacier—a tongue of the main ice-cap, Inglefield Gulf region, north Greenland, showing projection of the upper layers over the lower. The over-set of these layers is probably mainly due to the faster melting of the dirty ice below, and the fluting of the under surface of the projections is probably mainly due to the water from above, but the *débris* belt which is the primary cause, appears to have been formed by shearing. It is uncertain whether the shearing has immediately contributed anything to the projection.

FIG. 13.—Eastern edge of Bowdoin Glacier—a tongue of the main ice-cap, Inglefield Gulf region, north Greenland, showing contortion and shearing of the ice layers; movement from right to left. A ridge of rock juts out into the ice a little to the left of the part shown, and this is doubtless the occasion for the distortion and shearing.

FIG. 14.—Side view of the Gable Glacier—a terminal lobe of the main ice-cap of Greenland, Inglefield Gulf region, showing the taking on of basal load, and the development of shear-planes.

FIG. 15.—Striated surface of crystalline rock, near the end of Blasé Dale Glacier, Disco Island, Greenland.





